The Poisson distribution

Will Monroe
July 12, 2017

with materials by Mehran Sahami and Chris Piech
Announcements: Problem Set 2

Due today!

(Cell phone location sensing)
Announcements: Problem Set 3

To be released later today.

Due next Wednesday, 7/19, at 12:30pm (before class).

(election prediction)
Announcements: Midterm

Two weeks from yesterday:

Tuesday, July 25, 7:00-9:00pm

Tell me by the end of this week if you have a conflict!
Review: Basic distributions

Many types of random variables come up repeatedly. Known frequently-occurring distributions lets you do computations without deriving formulas from scratch.

\[ X \sim \text{Bin}(n, p) \]

We have ________ independent __________, \text{ INTEGER } \text{ PLURAL NOUN }

each of which ________ with probability \text{ VERB ENDING IN -S }

________. How many of the ________ \text{ REAL NUMBER } \text{ REPEAT PLURAL NOUN }

_______? \text{ REPEAT VERB -S }
An indicator variable (a possibly biased coin flip) obeys a **Bernoulli distribution**. Bernoulli random variables can be 0 or 1.

\[ X \sim \text{Ber}(p) \]

\[ p_X(1) = p \]
\[ p_X(0) = 1 - p \quad (0 \text{ elsewhere}) \]
Review: Bernoulli fact sheet

\[ X \sim \text{Ber}(p) \]

probability of “success” (heads, ad click, ...)

PMF:

\[ p_X(1) = p \]
\[ p_X(0) = 1 - p \quad \text{(0 elsewhere)} \]

expectation:

\[ E[X] = p \]

variance:

\[ \text{Var}(X) = p(1-p) \]

image (right): Gabriela Serrano
The **number of heads** on $n$ (possibly biased) coin flips obeys a **binomial distribution**.

$$X \sim \text{Bin}(n, p)$$

$$p_X(k) = \begin{cases} 
{n \choose k} p^k (1-p)^{n-k} & \text{if } k \in \mathbb{N}, 0 \leq k \leq n \\
0 & \text{otherwise}
\end{cases}$$
Review: Binomial fact sheet

number of trials (flips, program runs, ...)

probability of “success” (heads, crash, ...)

\[ X \sim \text{Bin}(n, p) \]

PMF:

\[
p_x(k) = \begin{cases} 
\binom{n}{k} p^k (1 - p)^{n-k} & \text{if } k \in \mathbb{N}, 0 \leq k \leq n \\
0 & \text{otherwise}
\end{cases}
\]

effectation:

\[ E[X] = np \]

Variance:

\[ \text{Var}(X) = np(1 - p) \]

Note: \[ \text{Ber}(p) = \text{Bin}(1, p) \]
The number of occurrences of an event that occurs with constant rate $\lambda$ (per unit time), in 1 unit of time, obeys a Poisson distribution.

$$X \sim \text{Poi}(\lambda)$$

$$p_X(k) = \begin{cases} e^{-\lambda} \frac{\lambda^k}{k!} & \text{if } x \in \mathbb{Z}, x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$
Algorithmic ride sharing

A ride-sharing program gets an average of 5 requests per minute from a region of the city.

X: number of requests this minute

Which looks like the most likely PMF for the distribution of X?


Room: CS109SUMMER17
Algorithmic ride sharing

A ride-sharing program gets an average of 5 requests per minute from a region of the city.

\( X \): number of requests this minute

Which looks like the most likely PMF for the distribution of \( X \)?


Room: CS109SUMMER17
Algorithmic ride sharing

A ride-sharing program gets an average of 5 requests per minute from a region of the city.

$X$: number of requests this minute

What is $P(X = 3)$?

1 minute = 0 1 0 0 1 ... 0 0 1 0 0

$X \sim \text{Bin}(n, p)$

image: Zbigniew Bzdak, Chicago Tribune
Algorithmic ride sharing

A ride-sharing program gets an average of 5 requests per minute from a region of the city.

$X$: number of requests this minute

What is $P(X = 3)$?

1 minute = $\begin{array}{cccccccc}
0 & 1 & 0 & 0 & 1 & \cdots & 0 & 0 & 1 & 0 & 0 \\
\end{array}$

$X \sim \text{Bin}(60, \frac{5}{60})$

$E[X] = 60 \quad p = \frac{5}{60}$

image: Zbigniew Bzdak, Chicago Tribune
Algorithmic ride sharing

A ride-sharing program gets an average of 5 requests per minute from a region of the city.

\(X: \text{number of requests this minute}\)

What is \(P(X = 3)\)?

1 minute = 60000 milliseconds

\[X \sim \text{Bin}(60000, \frac{5}{60000})\]

\[E[X] = 60000 \quad p = 5\]

\[p = \frac{5}{60000}\]
A ride-sharing program gets an **average** of 5 requests per minute from a region of the city.

Let $X$: number of requests **this** minute

What is $P(X = 3)$?

1 minute = $\infty$ instants

$X \sim \text{Bin}(\infty, \frac{5}{\infty})$

$E[X] = \infty$, $p = 5$

$p = \frac{5}{\infty}$
Binomial in the limit

\[ P(X = k) = \lim_{n \to \infty} \binom{n}{k} \left( \frac{5}{n} \right)^k \left( 1 - \frac{5}{n} \right)^{n-k} \]

\[ = \lim_{n \to \infty} \frac{n!}{k!(n-k)!} \frac{5^k (1-5/n)^n}{n^k (1-5/n)^k} \]

\[ = \lim_{n \to \infty} \frac{n(n-1)\cdots(n-k+1)}{k!} \frac{5^k (1-5/n)^n}{n^k (1-5/n)^k} \]

\[ = \lim_{n \to \infty} \frac{n^{n-1}}{n} \cdots \frac{n-k+1}{n} \frac{5^k (1-5/n)^n}{k! (1-5/n)^k} \]

\[ \frac{n}{n}, \frac{n-1}{n}, \ldots, \frac{n-k+1}{n} \to 1 \quad (1-5/n)^k \to 1^k = 1 \]
Binomial in the limit

\[ P(X=k) = \lim_{n \to \infty} \binom{n}{k} \left( \frac{5}{n} \right)^k \left( 1 - \frac{5}{n} \right)^{n-k} \]

\[ = \lim_{n \to \infty} \frac{n!}{k!(n-k)!} \frac{5^k}{n^k} \left( \frac{1 - 5/n}{1 - 5/n} \right)^k \]

\[ = \lim_{n \to \infty} \frac{n(n-1)\cdots(n-k+1)}{k!} \frac{5^k}{n^k} \left( \frac{1 - 5/n}{1 - 5/n} \right)^k \]

\[ = \lim_{n \to \infty} \frac{n}{n} \frac{n-1}{n} \cdots \frac{n-k+1}{n} \frac{5^k}{k!} \left( \frac{1 - 5/n}{1 - 5/n} \right)^k \]

\[ = \frac{5^k}{k!} e^{-5} \]
A ride-sharing program gets an average of 5 requests per minute from a region of the city.

**X**: number of requests this minute

What is $P(X = 3)$?

1 minute = $\infty$ instants

$$X \sim \text{Bin}(\infty, \frac{5}{\infty})$$

$$P(X = 3) = \frac{5^3}{3!} e^{-5} \approx 0.14$$
A ride-sharing program gets an average of 5 requests per minute from a region of the city.

$X$: number of requests this minute

What is $P(X = 3)$?

1 minute = 1 instant

$x \sim \text{Poi}(5)$

$P(X = 3) = \frac{5^3}{3!}e^{-5} \approx 0.14$
Poisson: Fact sheet

\( X \sim \text{Poi}(\lambda) \)

rate of events (requests, earthquakes, chocolate chips, ...)
per unit time (hour, year, cookie, ...)

PMF:
\[
p_X(k) = \begin{cases} 
  e^{-\lambda} \frac{\lambda^k}{k!} & \text{if } k \in \mathbb{Z}, k \geq 0 \\
  0 & \text{otherwise}
\end{cases}
\]
Make a very large chocolate chip cookie recipe.

Recipe tells you the overall ratio of chocolate chips per cookie ($\lambda$).

But some cookies get more, some get less!

$X \sim \text{Poi}(\lambda)$ is the number of chocolate chips in some individual cookie.

(This is called a “Poisson process”: independent discrete events [chocolate chips] scattered throughout continuous time or space [batter], with constant overall rate.)
Poisson: Fact sheet

$X \sim \text{Poi} (\lambda)$

rate of events (requests, earthquakes, chocolate chips, ...)
per unit time (hour, year, cookie, ...)

PMF:

$p_X(k) = \begin{cases} 
  e^{-\lambda} \frac{\lambda^k}{k!} & \text{if } k \in \mathbb{Z}, k \geq 0 \\
  0 & \text{otherwise}
\end{cases}$

expectation:

$E[X] = \lambda$
Expectation of Poisson

The cheater's way:

\[ \text{Poi}(\lambda) = \text{Bin}(n=\infty, p=\frac{\lambda}{\infty}) \]

\[ E[X] = np = \infty \cdot \frac{\lambda}{\infty} = \lambda \]

The real way:

\[ E[X] = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \cdot k \]

\[ = \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \cdot k \]

\[ = \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^k}{(k-1)!} \]

\[ = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \]

\[ = \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \]

\[ = \lambda e^{-\lambda} \cdot e^\lambda \]

\[ = \lambda \]

\[ \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = e^\lambda \]
Poisson: Fact sheet

\[ X \sim \text{Poi}(\lambda) \]

rate of events (requests, earthquakes, chocolate chips, ...) per unit time (hour, year, cookie, ...)

PMF:

\[ p_X(k) = \begin{cases} 
  e^{-\lambda} \frac{\lambda^k}{k!} & \text{if } k \in \mathbb{Z}, k \geq 0 \\
  0 & \text{otherwise} 
\end{cases} \]

expectation:

\[ E[X] = \lambda \]
Variance of Poisson

The cheater's way:

\[ \text{Poi}(\lambda) = \text{Bin}(n=\infty, p=\frac{\lambda}{\infty}) \]

\[ \text{Var}(X) = np(1-p) = \infty \cdot \frac{\lambda}{\infty} \cdot (1 - \frac{\lambda}{\infty}) = \lambda \cdot 1 = \lambda \]

The real way:

\[ E[X^2] = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \cdot k^2 = \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \cdot k^2 \]

\[ = \lambda \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} \cdot k \]

\[ = \lambda \sum_{j=0}^{\infty} e^{-\lambda} \frac{\lambda^j}{j!} \cdot (j+1) \]

\[ = \lambda \left[ \sum_{j=0}^{\infty} e^{-\lambda} \frac{\lambda^j}{j!} \cdot j + \sum_{j=0}^{\infty} e^{-\lambda} \frac{\lambda^j}{j!} \right] \]

\[ = \lambda \left[ E[X] + e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \right] \]

\[ = \lambda \left( \lambda + 1 \right) \]

\[ \text{Var}(X) = E[X^2] - (E[X])^2 = \lambda \left( \lambda + 1 \right) - \lambda^2 = \lambda \]
Poisson: Fact sheet

\( X \sim \text{Poi}(\lambda) \)

rate of events (requests, earthquakes, chocolate chips, ...)
per unit time (hour, year, cookie, ...)

PMF:
\[
p_X(k) = \begin{cases} 
  e^{-\lambda} \frac{\lambda^k}{k!} & \text{if } k \in \mathbb{Z}, k \geq 0 \\
  0 & \text{otherwise}
\end{cases}
\]

expectation:
\[ E[X] = \lambda \]

variance:
\[ \text{Var}(X) = \lambda \]
Siméon Denis Poisson

French mathematician (1781-1840)

- First paper at age 18
- Became professor at 21
- Published over 300 papers

« La vie n'est bonne qu'à deux choses : à faire des mathématiques et à les professer. »
—quoted in Arago (1854)

(“Life is only good for two things: doing mathematics and teaching it.”)
Break time!
Poisson: Fact sheet

\( X \sim \text{Poi}(\lambda) \)

rate of events (requests, earthquakes, chocolate chips, ...) per unit time (hour, year, cookie, ...)

PMF:

\[
p_x(k) = \begin{cases} 
  e^{-\lambda} \frac{\lambda^k}{k!} & \text{if } k \in \mathbb{Z}, k \geq 0 \\
  0 & \text{otherwise}
\end{cases}
\]

expectation:

\( E[ X ] = \lambda \)

variance:

\( \text{Var}( X ) = \lambda \)
Web server hits

Your web server gets an average of 2 requests each second.

What’s the probability of exactly 5 requests in a second?

**X**: number of requests in next sec.

\[ P(X = 5) = e^{-2} \frac{2^5}{5!} \approx 0.0361 \]
Earthquakes

There are an average of 2.8 major earthquakes in the world each year.

What’s the probability of >1 major earthquake next year?

\[ X: \text{number of earthquakes next year} \]

\[ P(X > 1) = 1 - P(X = 0) - P(X = 1) \]

\[ = 1 - e^{-2.8} \frac{2.8^0}{0!} - e^{-2.8} \frac{2.8^1}{1!} \]

\[ = 1 - 0.06 - 0.17 \]

\[ = 0.77 \]
IS THE SEQUENCE OF EARTHQUAKES IN SOUTHERN CALIFORNIA, WITH AFTERSHOCKS REMOVED, POISSONIAN?

By J. K. Gardner and L. Knopoff

Abstract

Yes.
We derived Poisson as the limit of the binomial as $n \to \infty$, with $\lambda = np$.

But it works (approximately) for finite $n$ as well!

$$\text{Bin}(n, p) \approx \text{Poi}(\lambda = np)$$

if: $n$ is large
$p$ is small
More space messages

Sending a 4000-bit message through space. Each bit corrupted (flipped) with probability \( p = 0.0001 \).

\[ X: \text{ number of bits flipped} \]
\[ X \sim \text{Bin}(4000, 0.0001) \]
\[ \approx \text{Poi}(0.4) \]

\[
P(X \leq 1) = P(X = 0) + P(X = 1)
\]
\[
= \binom{4000}{0} (0.0001)^0 (0.9999)^{4000-0} + \binom{4000}{1} (0.0001)^1 (0.9999)^{4000-1}
\]
\[
= 0.9999^{4000} + 4000 \cdot 0.0001 \cdot 0.9999^{3999}
\]
\[
\approx 0.67031 + 0.26815 = 0.93846
\]
Poisson approximation of binomial

$P(X = k)$

$k$

$k$ values: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Graph showing:
- $\text{Bin}(10, 0.3)$
- $\text{Bin}(100, 0.03)$
- $\text{Poi}(3)$
Poisson approximations are chill

Poisson often still works despite “mild” violations of the binomial distribution assumptions:

- Independence
  
  e.g., # of entries in each bucket in a hash table

- Same $p$:

  e.g. prob. of birthdays is somewhat uneven
Birthdays

$m$ people in a room.
What is the probability none share the same birthday?

\[ \binom{m}{2} \text{ trials, one for each distinct pair of people } (x, y) \]

\[ E_{xy} : x \text{ and } y \text{ have the same birthday ("success")} \]

\[ P(E_{xy}) \approx 1/365 \quad \text{not all independent} \]

Approximate: \( X \sim \text{Poi}(\lambda) \)

\[ \lambda = \left( \binom{m}{2} \right) \frac{1}{365} \]

\[ P(X = 0) = e^{-\left( \binom{m}{2} \frac{1}{365} \right)} \cdot \frac{\left( \binom{m}{2} \frac{1}{365} \right)^0}{0!} = e^{-\left( \binom{m}{2} \frac{1}{365} \right)} \]

\[ P(X = 0) < 1/2: n = 23 \]
A real license plate seen at Stanford
Randomness and clumping

uniformly random  “low-discrepancy” pseudorandom  regular pattern
Application: Low-discrepancy sampling for rendering

uniformly random

“low-discrepancy” pseudorandom

regular pattern
Philosophical note:
The structure of the universe

image: Andrew Z. Colvin