Poisson Process, Pokémons, and Prime Numbers

Kevin Wang

School of Mathematics and Statistics
The University of Sydney

Last Modified: October 27, 2016

Background
- That French guy

Poisson Process
- Definition

Coupon Collector's Problem revisited

Number Theory
- Prime Number Theorem and distribution of primes

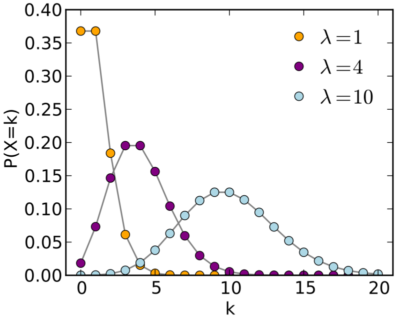
The Poisson Distribution

Named after Simeon Denis Poisson.

$$\mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \qquad k \in \mathbb{N}. \tag{1}$$

PP can be used to describe:

- radioactive decays,
- arrivals of buses
- or failures of Carslaw lift. (ongoing research)

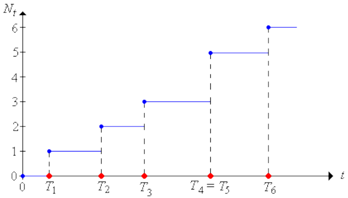


Counting process

Definition

A set of random variables $\{N(t)\}_{t\in\mathbb{N}}$ is called a **counting process** if:

1. $N(t)$ takes value in $\mathbb{N}$.
2. $N(0) = 0$.
3. If $s < t$, then $N(s) \leq N(t)$.
4. For $s < t$, $N(t) - N(s)$ equals the number of event in the interval $(s, t]$.



Poisson Process Definition

Definition

Furthermore, a Poisson Process has additional assumptions:

1. The random variable $N(s+t) - N(s)$ is $Pois(\lambda t)$ distributed, for **all** $s, t \geq 0$.
2. *Independent increments:* the numbers of events that occur in **disjoint** time intervals are independent.

- The time gap between any two consecutive arrivals (**interarrival time**) is $Exp(\lambda)$ distributed. $\mathbb{P}(X \geq t) = 1 - e^{-\lambda t}$.

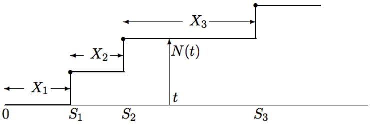


Figure 2.1: An arrival process and its arrival epochs $\{S_1, S_2, \ldots\}$, its interarrival intervals $\{X_1, X_2, \ldots\}$, and its counting process $\{N(t);\, t \geq 0\}$

Poisson Process Definition

Definition

Furthermore, a Poisson Process has additional assumptions:

1. The random variable $N(s+t) - N(s)$ is $Pois(\lambda t)$ distributed, for **all** $s, t \geq 0$.
2. *Independent increments:* the numbers of events that occur in **disjoint** time intervals are independent.

- The time gap between any two consecutive arrivals (**interarrival time**) is $Exp(\lambda)$ distributed. $\mathbb{P}(X \geq t) = 1 - e^{-\lambda t}$.

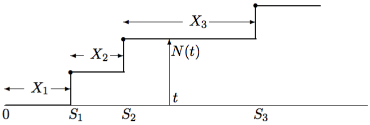


Figure 2.1: An arrival process and its arrival epochs $\{S_1, S_2, \ldots\}$, its interarrival intervals $\{X_1, X_2, \ldots\}$, and its counting process $\{N(t); t \geq 0\}$

Poisson Process Definition

Definition

Furthermore, a Poisson Process has additional assumptions:

1. The random variable $N(s+t) - N(s)$ is $Pois(\lambda t)$ distributed, for **all** $s, t \geq 0$.
2. *Independent increments:* the numbers of events that occur in **disjoint** time intervals are independent.

- The time gap between any two consecutive arrivals (**interarrival time**) is $Exp(\lambda)$ distributed. $\mathbb{P}(X \geq t) = 1 - e^{-\lambda t}$.

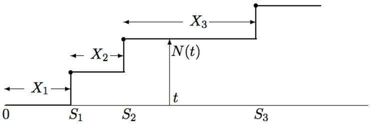


Figure 2.1: An arrival process and its arrival epochs $\{S_1, S_2, \dots\}$, its interarrival intervals $\{X_1, X_2, \dots\}$, and its counting process $\{N(t);\, t \geq 0\}$

The Coupon Collector's Problem with unequal probability

Problem

There are n individual coupons/Pokemons, each with capture probability p_i. Assuming we can collect 1 coupon per unit time, what is the expected number of coupon do we need to complete the collection?

- We also derived that in the unequal probability case:

$$\mathbb{E}(X) = \int_0^\infty \left(1 - \prod_{i=1}^{n} (1 - \exp(-p_i x))\right) \, dx. \tag{2}$$

- The Poisson Process solution is much cleaner.

Poisson Process Solution to CCP (Sketch)

- Modelling the **time** of arrival of coupons instead of **number** of coupons needed allow us to exploit *independence* of inter-arrival times.
- It can be shown that the *expected time for collecting all coupons* and *the number of coupons needed in total* are the **same** using properties of inter-arrival times of PP.

Poisson Process Solution to CCP (Sketch)

- Modelling the **time** of arrival of coupons instead of **number** of coupons needed allow us to exploit *independence* of inter-arrival times.
- It can be shown that the *expected time for collecting all coupons* and *the number of coupons needed in total* are the **same** using properties of inter-arrival times of PP.

Poisson Process Solution to CCP

- This meant $Z = \max\{Z_1, \ldots, Z_n\}$ is the maximum of a set of *independent* $Exp(p_i)$ random variable:

$$\begin{aligned} \mathbb{P}(Z \leq t) &= \mathbb{P}(Z_1 \leq t, \ldots, Z_n \leq t), \\ &= \prod_{i=1}^{n} \mathbb{P}(Z_i \leq t), \\ &= \prod_{i=1}^{n} (1 - \exp(-p_i t)) \, . \end{aligned} \tag{3}$$

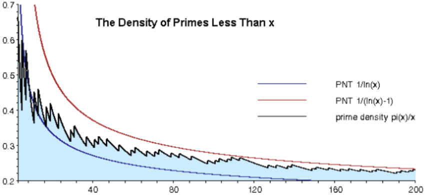
$$\begin{aligned} \mathbb{E}(Z) &= \int_0^\infty \mathbb{P}(Z > t) \\ &= \int_0^\infty \left(1 - \prod_{i=1}^{n} (1 - \exp(-p_i t)) \right) \, dt. \end{aligned} \tag{4}$$

Prime Number Theorem

- Define $\pi(x)$ as the function which counts the number of primes up to x.
- (One version of) the Prime Number Theorem states:

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log(x)} = 1. \tag{5}$$

- Gauss gaussed it (pun intended).
- Heuristically, since about $x/\log(x)$ of the x positive integers less than or equal to x are prime, the "probability" of one of them being prime is about $1/\log(x)$.

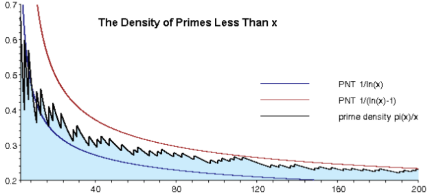


Prime Number Theorem

- Define $\pi(x)$ as the function which counts the number of primes up to x.
- (One version of) the Prime Number Theorem states:

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log(x)} = 1. \tag{5}$$

- Gauss gaussed it (pun intended).
- Heuristically, since about $x/\log(x)$ of the x positive integers less than or equal to x are prime, the "probability" of one of them being prime is about $1/\log(x)$.



Statistics of primes

- Harald Cramér (1893-1985) made some significant contributions/conjectures towards understanding the distribution of primes.
- "We are interested in the distribution of a given sequence S of integers, we then consider S as a member of an infinite class C of sequences, which may be concretely interpreted as the possible realizations of some game of chance. It is then in many cases possible to prove that, with a probability 1, a certain relation R holds in C, i.e. that in a definite mathematical sense 'almost all' sequences of C satisfy R". [1]
- This "pseudo-randomness" gives us some heuristic evidence that some conjectures are true.
- Cramér's Conjecture:

$$p_{n+1} - p_n = O((\log p_n)^2). \tag{6}$$

[1] "Of course we cannot in general conclude that R holds for the particular sequence S, but results suggested in this way may sometimes afterwards be rigorously proved by other methods".

Statistics of primes

- Harald Cramér (1893-1985) made some significant contributions/conjectures towards understanding the distribution of primes.
- "We are interested in the distribution of a given sequence S of integers, we then consider S as a member of an infinite class C of sequences, which may be concretely interpreted as the possible realizations of some game of chance. It is then in many cases possible to prove that, with a probability 1, a certain relation R holds in C, i.e. that in a definite mathematical sense 'almost all' sequences of C satisfy R". [1]
- This "pseudo-randomness" gives us some heuristic evidence that some conjectures are true.
- Cramér's Conjecture:

$$p_{n+1} - p_n = O((\log p_n)^2). \tag{6}$$

[1] "Of course we cannot in general conclude that R holds for the particular sequence S, but results suggested in this way may sometimes afterwards be rigorously proved by other methods".

Statistics of primes

- Harald Cramér (1893-1985) made some significant contributions/conjectures towards understanding the distribution of primes.
- "We are interested in the distribution of a given sequence S of integers, we then consider S as a member of an infinite class C of sequences, which may be concretely interpreted as the possible realizations of some game of chance. It is then in many cases possible to prove that, with a probability 1, a certain relation R holds in C, i.e. that in a definite mathematical sense 'almost all' sequences of C satisfy R". [1]
- This "pseudo-randomness" gives us some heuristic evidence that some conjectures are true.
- Cramér's Conjecture:

$$p_{n+1} - p_n = O((\log p_n)^2). \tag{6}$$

[1] "Of course we cannot in general conclude that R holds for the particular sequence S, but results suggested in this way may sometimes afterwards be rigorously proved by other methods".

Poisson distribution in primes

- Under Cramér's idea, we can imagine drawing 1 black/white ball every time from an infinite series urns, with probability of a white ball in urn U_n being $1/\log(n)$.
- Define

$$z_n := \mathbb{I}(n\text{-th urn gives a white ball}), \tag{7}$$

$$\Pi(x) := \sum_{n \leq x} z_n. \tag{8}$$

- Under Cramér's framework, $\Pi(x)$ is a random variable analogy of $\pi(x)$.
- Now, if n is close to some x, then we can think of $z_n \sim Binomial(1, 1/\log(x))$.
- For fixed $\lambda > 0$, $k \in \mathbb{N}$,

$$\#\{\text{integers } x \leq X : \Pi(x + \lambda \log(x) - \Pi(x) = k \sim \mathrm{Pois}(\lambda), \tag{9}$$

as $X \to \infty$.

Poisson distribution in primes

- Under Cramér's idea, we can imagine drawing 1 black/white ball every time from an infinite series urns, with probability of a white ball in urn U_n being $1/\log(n)$.
- Define

$$z_n := \mathbb{I}(n\text{-th urn gives a white ball}), \tag{7}$$

$$\Pi(x) := \sum_{n \leq x} z_n. \tag{8}$$

- Under Cramér's framework, $\Pi(x)$ is a random variable analogy of $\pi(x)$.
- Now, if n is close to some x, then we can think of $z_n \sim Binomial(1, 1/\log(x))$.
- For fixed $\lambda > 0$, $k \in \mathbb{N}$,

$$\#\{\text{integers } x \leq X : \Pi(x + \lambda \log(x) - \Pi(x) = k \sim \mathrm{Pois}(\lambda), \tag{9}$$

as $X \to \infty$.

Poisson distribution in primes

- Under Cramér's idea, we can imagine drawing 1 black/white ball every time from an infinite series urns, with probability of a white ball in urn U_n being $1/\log(n)$.
- Define

$$z_n := \mathbb{I}(n\text{-th urn gives a white ball}), \tag{7}$$

$$\Pi(x) := \sum_{n \leq x} z_n. \tag{8}$$

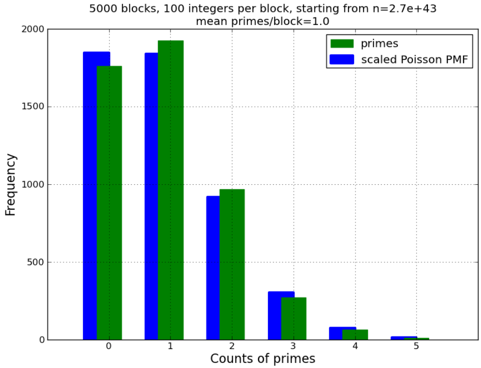
- Under Cramér's framework, $\Pi(x)$ is a random variable analogy of $\pi(x)$.
- Now, if n is close to some x, then we can think of $z_n \sim Binomial(1, 1/\log(x))$.
- For fixed $\lambda > 0$, $k \in \mathbb{N}$,

$$\#\{\text{integers } x \leq X : \Pi(x + \lambda \log(x) - \Pi(x) = k \sim \mathrm{Pois}(\lambda), \tag{9}$$

as $X \to \infty$.

Poisson process and primes

- In other words, under this pseudo-random model of the primes, the random sets "behave" like a Poisson process:



References

- Robert Gallager. Lecture notes on Poisson Process.
- http://phillipmfeldman.org/mathematics/primes.html
- https://www.youtube.com/watch?v=pp06oGD4m00
- *Harald Cramér and the distribution of prime numbers* A. Granville. 1993.
- *The Coupon Collector's Problem*, M. Ferrante, M. Saltalamacchia, (2014)
- *Introduction to Probability Models.* S.Ross.
- *A First Course in Probability.* S. Ross.
- *STAT3911: Stochastic Processes Lecture Notes.* R. Kawaii.
- https://primes.utm.edu/howmany.html
- https://terrytao.wordpress.com/tag/cramers-random-model/