MATH 180A: INTRO TO PROBABILITY
(FOR DATA SCIENCE)

www.math.ucsd.edu/~tkemp/180A

Today: § 4.6

Next: § 5.1 - 5.2

Lab4 due TONIGHT by 11:59pm

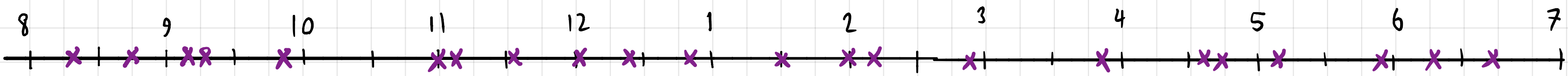
HW 5 due Friday (Nov 8) by 11:59pm

→ Error in problem statement of Exercise 4.40: the exact probability is $\doteq 0.00327556$

$\neq 0.00094968$

Li-Tien the lemur hangs out at Turtle Rock all day. He observes that cars pass by rarely, randomly, on average every 30 minutes.

He marks the times they come by on a number line:



4 in $(9,11]$

8 in $(12,16]$

8 in $(10,14]$

Over any time interval $(a,b]$, "on average" there are $2(b-a)$ cars.

Model: $N(a,b] = \#\text{cars in } (a,b] \sim \text{Poisson}(2(b-a))$

Poisson Process

A Poisson process of rate $\lambda > 0$ is a collection of distinct random points in $[0,\infty)$ with the following properties:

(1) For any bounded interval $(a,b]$ $(0 \le a < b < \infty)$ the number

$$N((a,b]) := \#\{\text{points in } (a,b]\} \sim \text{Poisson}(\lambda(b-a))$$

(2) For any non-overlapping intervals $I_1 = (a_1,b_1], \dots, I_k = (a_k,b_k]$, the random variables $N(I_1), N(I_2), \dots, N(I_k)$ are <u>independent</u>.

Useful notation: $N_t = N((0,t])$. Then $N((a,b]) = N_b - N_a$

"increments"

(1) $N_b - N_a \sim \text{Poisson}(\lambda(b-a))$

(2) $0 \le a_1 < b_1 \le a_2 < b_2 \le \cdots \le a_k < b_k < \infty$

$N_{b_1} - N_{a_1}, \; N_{b_2} - N_{a_2}, \; \dots, \; N_{b_k} - N_{a_k}$ are independent

$(N_t)_{t \geq 0}$ is an example of a stochastic process.

family of random variables indexed by time; usually with nice properties relating the distributions at different times. (MATH 180B/C)

Special kind of stochastic process, with independent increments. (Sometimes called a Lévy process.)

Not easy to achieve.

Fix $0 < s < t < \infty$.

$$N_t = N_s + N_t - N_s$$

$N_t \sim \text{Poisson}(\lambda t)$

$N_s = X \sim \text{Poisson}(\lambda s)$

$N_t - N_s = Y \sim \text{Poisson}(\lambda(t-s))$

$$N_t = X + Y$$

independent

Theorem: Let $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\mu)$.
If X, Y are independent, then

$$X + Y \sim \text{Poisson}(\lambda + \mu)$$

Proof: Think about bionomials (of the same success rate).

$$\text{indep} \begin{cases} S_n \sim \text{Bin}(n,p) & S_n = X_1 + X_2 + \cdots + X_n, \quad X_j \text{ ind. Ber}(p) \\ T_m \sim \text{Bin}(m,p) & T_m = Y_1 + Y_2 + \cdots + Y_m, \quad Y_j \text{ ind. Ber}(p) \end{cases}$$

$$S_n + T_m = \underbrace{X_1 + X_2 + \cdots + X_n + Y_1 + Y_2 + \cdots + Y_m}_{n+m \text{ ind. Ber}(p)} \sim \text{Bin}(n+m, p)$$

$$\left.\begin{aligned} X &\sim \text{Poiss}(\lambda) \approx \text{Bin}\left(n, \tfrac{\lambda}{n}\right) \sim S_n \\ Y &\sim \text{Poiss}(\mu) \approx \text{Bin}\left(m, \tfrac{\mu}{m}\right) \sim T_m \end{aligned}\right\} X + Y \approx S_n + T_m \sim \text{Bin}(n+m, p)$$

$$\approx \text{Poisson}((n+m)p)$$

$$= \text{Poisson}(\lambda + \mu) \quad ///$$

$$\frac{\mu}{m} = \frac{\lambda}{n} = p \implies (n+m)p = np + mp = \lambda + \mu$$

Example: Customers arriving at the Art of Espresso in the afternoon arrive according to a Poisson process with rate 20/hour.

(a) What is the probability no one comes between 2pm and 2:15pm?

$$N(s,t] \sim \text{Poisson}(20(t-s)) \sim \text{Poisson}(5) \qquad \mathbb{P}(N=k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

$s = 2$, $t = 2\frac{1}{4}$

$$\mathbb{P}(N_{2\frac{1}{4}} - N_2 = 0) = e^{-5} \doteq 0.674\%$$

(b) What is the probability that 2 customers come between 2pm and 2:15pm, and 5 customers come between 2:15pm and 2:20pm?

$$\mathbb{P}(N_{2\frac{1}{4}} - N_2 = 2 \ \& \ N_{2\frac{1}{3}} - N_{2\frac{1}{4}} = 5)$$

$(2, 2\frac{1}{4}] \leftarrow \rightarrow (2\frac{1}{4}, 2\frac{1}{3}]$ non-overlapping.

$$= \mathbb{P}(\underbrace{N_{2\frac{1}{4}} - N_2}_{\substack{\text{Poisson}(20\cdot\frac{1}{4}) \\ \text{Poisson}(5)}} = 2)\,\mathbb{P}(\underbrace{N_{2\frac{1}{3}} - N_{2\frac{1}{4}}}_{\substack{\text{Poisson}(20\cdot(\frac{1}{3}-\frac{1}{4})) \\ \text{Poisson}(5/3)}} = 5) = \left(e^{-5}\cdot\frac{5^2}{2!}\right)\left(\frac{e^{-\frac{5}{3}}\left(\frac{5}{3}\right)^5}{5!}\right)$$

$$\doteq 0.1704\%$$

Let $(N_t)_{t\geq 0}$ be a Poisson process of intensity $\lambda > 0$.

Let T be the time of the first jump (i.e. first customer/car/etc.)

$$N_t = \begin{cases} 0 & t<T \\ 1 & t=T \end{cases}$$

T is random.

What is the distribution of T? I.e. find its CDF.

$$F_T(t) = \underset{\substack{\parallel \\ 0 \text{ if } t<0.}}{\mathbb{P}(T \leq t)} = 1 - \mathbb{P}(T>t)$$

$\{T>t\} = \{\text{no calls in } (0,t]\}$

$$\mathbb{P}(T>t) = \underbrace{\mathbb{P}(N_t = 0)}_{\text{Poisson}(\lambda t)} = e^{-\lambda t}$$

$$= 1 - e^{-\lambda t}$$

$T \sim \text{Exp}(\lambda)$

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & t<0. \end{cases}$$

Constant Renewal

Let $(N_t)_{t\geq 0}$ be a Poisson process of intensity $\lambda > 0$.

Fix some $t_0 \geq 0$.

Define a new process $M_t = N_{t_0+t} - N_{t_0}$. What kind of process is it?

↑

is a Poisson(λ).

$$M_b - M_a = (N_{b+t_0} - \cancel{N_{t_0}}) - (N_{a+t_0} - \cancel{N_{t_0}})$$

$$= N_{b+t_0} - N_{a+t_0} \sim \text{Poisson}\Big(\lambda\big(\underbrace{(b+t_0) - (a+t_0)}_{=b-a}\big)\Big)$$

Has the right increments.

Also independent (similar argument). $\Big\}$ M_t is a Poisson process of intensity λ.

Example: In a Poisson process of intensity λ, what is the distribution of the time interval between the 999^{th} event and the 1000^{th}?

Let t_{999} = time of the 999^{th} event. Then $M_t = N_{t+t_{999}} - N_{t_{999}}$ is a Poisson(λ) process. Let T be the time of the 1000^{th} event for N_t; this is the time of the $\underline{1^{st}}$ event for M_t!

So $T \sim \text{Exp}(\lambda)$.

How about T_2 = time of second jump?

(Come back to this later)