Poisson

18.600 Problem Set 5, due March 24

Welcome to your fifth 18.600 problem set! We'll be thinking more about Poisson random variables and the corresponding processes. Please stop by my weekly office hours (2-249, Wednesday 3 to 5) for discussion.

A. FROM TEXTBOOK CHAPTER FOUR:

1. Theoretical Exercise 16: Let X be a Poisson random variable with parameter λ. Show that $P\{X = i\}$ increases monotonically and then decreases monotonically as i increases, reaching its maximum when i is the largest integer not exceeding λ. *Hint:* Consider $P\{X = i\}/P\{X = i - 1\}$.

2. Theoretical Exercise 25: Suppose that the number of events that occur in a specified time is a Poisson random variable with parameter λ. If each event is "counted" with probability p, independently of every other event, show that the number of events that are counted is a Poisson random variable with parameter λp. Also, give an intuitive argument as to why this should be so. As an application of the preceding result, suppose that the number of distinct uranium deposits in a given area is a Poisson random variable with parameter $\lambda = 10$. If, in a fixed period of time, each deposit is discovered independently with probability $\frac{1}{50}$, find the probability that (a) exactly 1, (b) at least 1, and (c) at most 1 deposit is discovered during that time.

B. ANSWER THE FOLLOWING:

1. Compute the expectation of X^n where n is a positive integer and X is a uniform random variable on the interval $[0, 1]$.

2. How does the answer change if the random variable is instead taken to be uniform on $[0, L]$ for some constant L?

C. In Regular Bus City, there is a shuttle bus that goes between Stop A and Stop B, with no stops in between. The bus is perfectly punctual and arrives at Stop A at precise five minute intervals (6:00, 6:05, 6:10, 6:15, etc.) day and night, at which point it immediately picks up all passengers waiting. Citizens of Regular Bus City arrive at Stop A at Poisson random times, with an average of 5 passengers arriving every minute, and board the next bus that arrives.

(a) Suppose that you visit this city and that you arrive at Stop A at a time chosen uniformly at random from the times in a day. How long do you expect to have to wait until the next bus?

(b) How many citizens of Regular Bus City do you expect to be on the bus that you take?

In Poisson Bus City, there is a shuttle bus that goes between Stop A and Stop B, with no stops in between. The times at which the bus arrives at Stop A are a Poisson point process with one bus arriving every five minutes on average, day and night, at which point it immediately picks up all passengers waiting. Citizens of Poisson Bus City (like those of Regular Bus City) arrive at Stop A at Poisson random times, with an average of 5 passengers arriving every minute, and board the next bus that arrives.

(c) Suppose that you visit this city and that you arrive at Stop A at a time chosen uniformly at random from the times in a day. How long do you expect to have to wait until the next bus?

(d) How many citizens of Poisson Bus City do you expect to be on the bus that you take?

(e) Are the following two statements true or false? If they are both true, explain in words the apparent discrepancy:

  (i) When you visit, buses in Poisson Bus City seem on average to come twice as slowly and to be twice as crowded as those in Regular Bus City

  (ii) In both cities, buses come on average every five minutes and people come on average five times per minutes, so that over the long haul there are 25 people per bus on average—so buses are on average equally crowded in the two cities.

Remark: Poisson Bus City is not the worst case scenario. Suppose that buses come in pairs (one right behind the other) with the pairs arriving as a Poisson point process with one pair every 10 minutes on average. And suppose that whenever this happens, everybody gets in the first bus and leaves the second bus empty. Now if you arrive at a random time, you can expect your bus to take four times as long to come and be four times as crowded as in Regular Bus City (assuming that like others you get on the first bus in a pair). On a real life bus route with many stops, the closer a bus is to the bus ahead of it, the faster it can go (since it is picking up fewer passengers) which can lead this kind of clumping.

D. Each day (independently of all other days) Jill has a one in five thousand chance of hearing the basic details of the Peleponnesian War. She stores something in long term memory after hearing it 4 times. Use Poisson approximations to (approximately) answer the following:

(a) What is the probability that, by the time Jill is 10,000 days old, she has the Peleponnesian War in long term memory?

Alice is more of a reader than Jill and also has a better memory for trivia. Each day (independently of all other days) Alice has a one in one thousand chance of learning about the Peleponnesian War, and she stores the information in long term memory after hearing it 3 times.

(b) What is the probability that, by the time Alice is 10,000 days old, she has the Peleponnesian War in long term memory?

(c) If there are $10,000$ similar facts (each fact comes with same probabilities as above), how many of them do we expect that Jill knows but Alice doesn't (assuming that both are 10,000 days old)? Assume that for each given fact, the two Poisson random variables (number of times fact is heard by Alice and by Jill) are independent. (If the answer is small, then Jill should feel pretty lucky when one of these facts comes up while she is watching Jeopardy with Alice.)

E. This problem addresses the Gompertz model for the duration of human life. But it starts out as another story about buses. Let X_1, X_2, X_3 be a Poisson point process of parameter 1 on $[0, \infty)$. Recall that this implies that X_1 and $X_2 - X_1$ and $X_3 - X_2$, etc., are i.i.d. exponential random variables each with parameter 1. Now for each integer $i \geq 1$, let $Y_i = \log X_i$. In Poisson Bus City, you might imagine that a bus line starts operating at time zero, and thereafter bus arrivals correspond to the times X_i.

On Accelerating Frequency Planet (AFP) the bus arrival times are $Y_1, Y_2, \ldots$. That is, each arrivial time is the *natural logarithm* of a point in the Poisson point process. Time is measured from $-\infty$ to ∞ on AFP, so it is possible that some bus arrival times are negative.

(a) Show that, on AFP, given any constants $a < b$, the number of buses that arrive between times a and b is a Poisson random variable with parameter $\int_a^b e^x dx$.

(b) Explain (with a sentence each) why the following things are true: the number of buses that arrive during the time interval $(-\infty, 0]$ is Poisson with parameter 1, while with probability one *infinitely* many buses arrive after time 0. If $\alpha = \ln 2 \approx .7$, then for each k the expected number of buses that arrive during $[k\alpha, (k+1)\alpha]$ is twice as large as the the expected number that arrive during $[(k-1)\alpha, k\alpha]$. (In other words, the doubling time for the bus arrival-frequency rate is α.) Moreover, the *first* bus's arrival time is a random variable whose median is $\ln\big(\ln(2)\big) \approx -.37$.

Note you can approximate an ordinary Poisson point process with parameter λ by partitioning time into disjoint intervals of the form $[t, t+\epsilon)$ for small ϵ and asserting that each interval independently contains a bus with probability $\lambda\epsilon$. Things are similar on AFP, except that the probability is approximately $e^t\epsilon$ when ϵ is small; in some sense, this is like saying that the Poisson parameter λ is "time dependent" (and exponentially increasing) with $\lambda(t) = e^t$.

More to the story: at time -7 on AFP, an adorable but immobile sloth is born at the bus stop, where it lives until it is killed by the first bus that arrives. Since it is unlikely the first bus comes before time -7, (b) suggests that the sloth's life span is a random variable with median about $7 - .37 \approx 6.63$. The standard unit of time on AFP is the duodecennium (i.e., twelve years), so that α "units" means $12\alpha \approx 8.32$ years and the sloth's median life span is $6.63 * 12 \approx 80$ years.

(c) Suppose that on AFP, half of the buses have fat tires and half have thin tires (bus type decided by independent coin toss for each bus), and female sloths are only killed by fat tired buses, while males are killed by all buses. Argue that if the sloth is female, its life expectancy is about 8.32 years (i.e., α duodecennia) longer than if it had been male. (Hint: use problem A.2 and argue that the probability density function for the lifespan of a female born at time -7 agrees with that of a male born at time $-7-\alpha$. Then note that having a bus between time $-7-\alpha$ and -7 is very unlikely.)

(d) Let p_k be the probability that the sloth dies during its kth year of life, *given* that it has survived for $(k-1)$ years. Argue that p_k is approximately the expected number of buses that arrive during that kth year when p_k is small — say, less than .1. (This is related to arguing that the probability that a λ Poisson random variable equals 1 is approximately λ when λ is reasonably small.) Thus p_k grows (roughly) exponentially in k for the first 80 or so years of life.

(e) Look up `https://www.ssa.gov/oact/STATS/table4c6.html` and `https://en.wikipedia.org/wiki/Gompertz%E2%80%93Makeham_law_of_mortality` and (after looking them over for five minutes) write a sentence or two about what you noticed — and in particular about how closely the p_k corresponding to humans match those of the sloths on AFP. (Three obvious differences: humans are much more likely than AFP sloths to die during the first year or so of life. Humans in late teens and twenties — especially males — die at a rate that is higher than the Gompertz law would predict, hence the "bump" in the otherwise nearly straight line on the Wikipedia chart. This may be in part be due to risky behaviors, which are either more pronounced among people that age or simply make a larger difference on the log scale because other causes of death are low. Finally, outside the "bump" period the gap betwen death rates for male and female humans is large but not as large as for AFP sloths.) You may find it helpful to consult `https://gravityandlevity.wordpress.com/2009/07/08/your-body-wasnt-built-to-last-a-lesson-from-human-mortality-rates/` for a somewhat breezier account of the Gompertz law story. (Notice also that this problem has another part after the next few remarks.)

REMARK: Gompertz law (i.e., exponential mortality rate growth) appears to apply pretty well to both humans and animals (with a species dependent doubling rate α). Google Gompertz morality and find out more. If you wanted a story to explain the exponential growth, one naive one would be that "glitches" in the body accumulate exponentially, and your death rate is proportional to the number of glitches. Another simplistic story is that if there is a genetic mutation that *causes* an organism to die at age X, then the rate at which natural selection eliminates that gene decreases with the proportion of organisms who reach age X. So mutations that cause functioning to break down in old people accumulate faster than those that affect young people. Note also that although Gompertz law holds pretty well in developed nations, it does not hold in settings where a large fraction of deaths are caused by predators, wars, infectious diseases, etc. that are as deadly to the young as the old.

REMARK: The assumption of an exponential increase in "bus arrival frequency" suggests that an unhealthy habit that *doubles* your probability of dying within any given year should *subtract* about α units of life expectancy (one doubling period). Medical advances that eliminate half the number of fatal buses should *add* about α units of life expectancy. If medicine eliminated 7/8 of the buses, this should increase life expectancy by 3α units, about 25 years. Lifestyle choices that increase death rate during any given year by 41 percent (i.e., which multiply death rate during any given year by $\sqrt{2}$) should take about $\alpha/2$ units (about four years) off life expectancy (since your death rate at any time is what it would be if you were 4 years older).

REMARK: In a world where Gompertz applies precisely with a doubling period of 8 years (and people's life spans are independent of each other), if you have $8 = 2^3$ students of age 20 and one professor of age $44 = 20 + 3 \times 8$, then the time until the first student dies should agree in law with the time until the professor dies. Essentially, being three doubling periods older means you expect eight times as many buses coming your way. But being eight people instead of one means that (as a group) you also have eight times as many buses coming your way. (If you want to extend the bus metaphor, imagine a different independent lane of buses for each person, with frequency rates depending on that person's age...) If you have a class of 128 students of age 20 and one older professor of age 76, then the time until the first student dies agrees in law with the time until the professor dies. If two parents are 24 years older than a child then, of the buses coming towards these three people, 1/17 are coming toward the child and 8/17 are coming toward each parent (since each parent has 8 times as many buses coming its way as the child). In fact, one can show that *given* the arrival times of the buses, we can assign each bus independently to one of the three people, with probabilities 8/17 for each parent and 1/17 for child. Looking at the first bus, we find that the probability that the child is the first of the three people to die is 1/17.

(f) Suppose that, in the world of the previous remark, there are three siblings of ages 20, 28, and 36. What is the probability the oldest one dies first?

REMARK: In order to *drastically* increase human life span (so we live to 150, say) we'll have to find a way to change α — i.e., to slow the exponential growth of the mortality rate. How can we do this? Can we freeze blood and tissue from when we're younger and reintroduce it later? Can we reset age markers? Can we manually edit the genes that cause aging? Somebody in this class should figure this out. We have a lot of buses coming our way.