STAT 150 HOMEWORK #8

FALL 2023

Due Friday, Oct 27th, at 11:59 PM on Gradescope.

Note that there are *Exercises* and *Problems* in Pinsky and Karlin. Make sure you read the homework carefully to find the assigned question.

1. Durrett, Exercise 2.36

Traffic on Snyder Hill Road in Ithaca, NY, follows a Poisson process with rate 2/3's of a vehicle per minute. 10 % of the vehicles are trucks, the other 90 % are cars.
(a) What is the probability at least one truck passes in a hour?
(b) Given that ten trucks have passed by in an hour, what is the expected number of vehicles that have passed by.
(c) Given that 50 vehicles have passed by in a hour, what is the probability there were exactly 5 trucks and 45 cars.

SOLUTION (A)

Pr(At least one Truck in One Hour) $= 1 - Pr$(Zero trucks in 60 Minutes) $=$
Because the rate of vehicles is $PP(\lambda = \frac{2}{3})$
then the rate of trucks is a simple thinning which is $PP(\lambda = \frac{2}{3} \cdot \frac{1}{10} = \frac{1}{15})$

Pr(Zero trucks in 60 Minutes) $= \frac{(\frac{1}{15} \cdot 60)^{(0)} e^{-(\frac{1}{15}) \cdot 60}}{(0)!} = e^{-(\frac{1}{15}) \cdot 60}$

Pr(At least one Truck in 60 Minutes) $= \boxed{1 - e^{-4}}$

SOLUTION (B)

Thinned PP's are independent. Thus, knowledge about the number of trucks does not affect our knowledge about the number of cars. Arrival of cars are distributed by $PP(\lambda = \frac{2}{3} \cdot \frac{9}{10})$ Our expectation for number of cars in one hours is: $60 \cdot \lambda$

$60 \cdot \mathbb{E}(\text{Cars in 1 Minute}) = 60 \cdot \frac{2}{3} \cdot \frac{9}{10} = 36$ Cars
$\boxed{36 \text{ Cars} + 10 \text{ Trucks} = 46 \text{ Vehicles}}$

SOLUTION (C)

As seen in PSET 6, a Poisson Process conditional on a number of events in binomial. Although the details are a bit different (two possible time periods vs two possible variable types) they fundamentally change the distribution in the same way (N Bernoulli trials where N is constant).

This makes sense, because the only variability will come from ability to flip the correct amount of each, and the number of possible arrangements. The rate λ won't affect this process of 'coin flipping' because it just affects the 'speed at which coins are flipped'

$$\boxed{Pr(\text{Exactly 5 cars and 45 trucks}) = \binom{50}{5}(\frac{1}{10})^{45}(\frac{9}{10})^{5}}$$

2. Durrett, Exercise 2.59

Signals are transmitted according to a Poisson process with rate λ. Each signal is successfully transmitted with probability p and lost with probability $1-p$. The fates of different signals are independent. For $t \geq 0$ let $N_1(t)$ be the number of signals successfully transmitted and let $N_2(t)$ be the number that are lost up to time t.

(a) Find the distribution of $(N_1(t), N_2(t))$.
(b) What is the distribution of $L =$ the number of signals lost before the first one is successfully transmitted?

SOLUTION (A):

The Poisson thinning theorem says that: $N_1(t) \sim Poisson(\lambda p)$ and $N_2(t) \sim Poisson(\lambda(1-p))$ and that they are both independent. The joint distribution of two independent Poisson's are just a product of the marginal distributions.

$$f_{(X,Y)}(x,y) = f_X(x) \cdot f_Y(y) = \frac{((\lambda p)t)^{n_1} \cdot e^{-(\lambda p)t} \cdot (\lambda(1-p)t)^{n_2} \cdot e^{-(\lambda(1-p))t}}{n_1! n_2!} = \frac{((\lambda p)t)^{n_1} \cdot (\lambda(1-p)t)^{n_2} \cdot e^{-(\lambda t)}}{n_1! n_2!}$$

The discrete joint PMF is: $\boxed{e^{-(\lambda t)} \cdot \dfrac{(\lambda t)^{n_1+n_2}(p)^{n_1}(1-p)^{n_2}}{n_1! n_2!}}$

SOLUTION(B):

Because each signal's transmission is determined by a Bernoulli(p) trial, and that each trial is independent, we are essentially waiting for the first success among failures. We know this is:

$P(L = k) = (1-p)k(p) \rightarrow \boxed{L \sim Geometric(p)}$.

Once again, notice how the rate λ does not show up in the final answer. This is similar to Problem 1C in that the variable in question is not a function of 'overall speed of flipping coins' but rather only interested in the outcomes of the coins in a specific arrangement.

3. Durrett, Exercise 2.61
Consider two independent Poisson processes $N_1(t)$ and $N_2(t)$ with rates λ_1 and λ_2. What is the probability that the two-dimensional process $(N_1(t), N_2(t))$ ever visits the point (i, j)?

SOLUTION:

The counts in any PP are monotonically increasing. $N_1(t)$ and $N_2(t)$ are both monotonically increasing. Notice then that the superposition of both processes, which I will denote $N(t)$ which is $Poisson(\lambda_1+\lambda_2)$ will eventually reach $(i+j)$ in finite time. At this time. $N_1(t)$ and $N_2(t)$ will account for given portions of the counts. We want to see, that given $N(t) = (i+j)$ what is the probability that $N_1(t) = i$ and $N_2(t) = j$.

Due to the Poisson thinning theorem, the probability that any occurrence of $N(t)$ has probability $\frac{\lambda_1}{\lambda_1+\lambda_2}$ of belonging to $N_1(t)$ and probability of $\frac{\lambda_2}{\lambda_1+\lambda_2}$ of belonging to $N_2(t)$.

We do not care about a specific arrangement of the arrivals and so we multiply by $\binom{i+j}{i}$ to denote all the possible ways we can arrange i items in $i+j$ spots. Thus we have reduce the problem to an 'arrangement of Bernoulli's' setup, where the Bernoulli trials have $p = \frac{\lambda_1}{\lambda_1+\lambda_2}$ which we know has binomial probabilities.

$$\boxed{Pr(N_1(t) = i, N_2(t) = j) = \binom{i+j}{i}(\frac{\lambda_1}{\lambda_1+\lambda_2})^{(i)}(\frac{\lambda_2}{\lambda_1+\lambda_2})^{(j)}}$$

(CAUTION FOR GRADER, IGNORE FOR GRADING PURPOSES) The bottom part is an original attempt which has conceptual mistakes. Taking the integral across time answers the question what is the expected amount of time we spend at (i, j). This is because, after reaching (i,j) we are not forced to leave in a set time period, and the time spent is random. However, we just want the static probability of EVER reaching that point. These are different questions.

Take the joint PMF. They are independent so just take the product of the marginals: $e^{-t(\lambda_1+\lambda_2)} \cdot \frac{(\lambda_1 t)^i(\lambda_2 t)^j}{i!j!}$

We need to integrate this across time, because we are interested in whether they will ever equal i and j.

$\int_0^\infty e^{-t(\lambda_1+\lambda_2)} \cdot \frac{(\lambda_1 t)^i(\lambda_2 t)^j}{i!j!} dt = \frac{(\lambda_1)^i(\lambda_2)^j}{i!j!}\int_0^\infty (t)^{i+j}e^{-t(\lambda_1+\lambda_2)}dt$

Substitute: $u = t(\lambda_1+\lambda_2)$, $du = dt(\lambda_1+\lambda_2)$

$\frac{(\lambda_1)^i(\lambda_2)^j}{i!j!}\int_0^\infty (\frac{u}{(\lambda_1+\lambda_2)})^{i+j}e^{-u}\frac{du}{(\lambda_1+\lambda_2)} = \frac{(\lambda_1)^i(\lambda_2)^j}{i!j!(\lambda_1+\lambda_2)^{i+j+1}}\int_0^\infty u^{i+j}e^{-u}du$
$\frac{(\lambda_1)^i(\lambda_2)^j}{i!j!(\lambda_1+\lambda_2)^{i+j+1}}\Gamma(i+j+1) = \frac{(\lambda_1)^i(\lambda_2)^j}{i!j!(\lambda_1+\lambda_2)^{i+j+1}}(i+j)! = \frac{(\lambda_1)^i(\lambda_2)^j}{(\lambda_1+\lambda_2)^{i+j+1}}\binom{i+j}{i}$
$Pr(N_1(t) = i, N_2(t) = j) = \binom{i+j}{i}\frac{1}{\lambda_1+\lambda_2}(\frac{\lambda_1}{\lambda_1+\lambda_2})^{(i)}(\frac{\lambda_2}{\lambda_1+\lambda_2})^{(j)}$

4. Pinsky and Karlin, EXERCISE 7.1.3

Which of the following are true statements?
(a) $N(t) < k$ if and only if $W_k > t$.
(b) $N(t) \leq k$ if and only if $W_k \geq t$.
(c) $N(t) > k$ if and only if $W_k < t$.

SOLUTION:

(a) TRUE
If there have been less than k arrivals by time t, then the waiting time for the k-th arrival must be greater than t. If the k-th arrival arrives after time t, then it is necessarily true that the there are less than k arrivals at time t.

(b) FALSE
Counterexample: If the k-th arival happens at $t-\delta$, then $W_k < t$. However, $N(t) = k$ is still satisfied, so long as there are no new arrivals until time t
(During the interval $(t, t + delta)$).

(c) FASLSE
Counterexample: The $k-th$ arrivals has occured before t. Thus, $W_k < t$. If, during the next $t - W_k$ amount of time, there are no new arrivals. Then $N(t) = k$ and thus we have disproved the statement.

5. Pinsky and Karlin, EXERCISE 7.3.3

Let $W_1, W_2, \ldots$ be the event times in a Poisson process $\{N(t); t \geq 0\}$ of rate λ. Show that $N(t)$ and $W_{N(t)+1}$ are independent by calculating:
$Pr\{N(t) = n)$ and $W_{N(t)+1} > t + s\}$

SOLUTION:

Preliminary: $(N(t) = n) = (W_n \leq t < W_{n+1})$

$Pr\{(N(t) = n)$ and $W_{N(t)+1} > t + s\} =$
$Pr\{(N(t) = n)$ and $W_{n+1} > s + t\} =$
$Pr\{(N(t) = n)$ and $W_{n+1} > s + t\} =$
$Pr\{(W_n \leq t < W_{n+1})$ and $W_{n+1} > s + t\} =$
$Pr\{(W_n \leq t)$ and $(t + s < W_{n+1})\} =$
$Pr\{(W_n \leq t)$ and $(t + s < W_n + X)\} =$
$Pr\{(W_n \leq t)$ and $(t + s - W_n < X)\} =$

Condition and integrate on $W_n = w$:

$\int_0^\infty Pr(((W_n \leq t)$ and $(t + s - W_n < X))|W_n = w)Pr\{W_n = w\}dw =$

Where $Pr\{W_n = w\}$ is $Gamma(r, \lambda)$

$\int_0^\infty Pr(((W_n \leq t)$ and $(t + s - W_n < X))|W_n = w)\frac{\lambda^n}{(n-1)!}w^{n-1}e^{-\lambda w}dw =$
$\int_0^\infty Pr(((w \leq t)$ and $(t + s - w < X))|W_n = w)\frac{\lambda^n}{(n-1)!}w^{n-1}e^{-\lambda w}dw =$
$\int_0^\infty (\mathbb{1}(w \leq t)Pr(t + s - w < X))\frac{\lambda^n}{(n-1)!}w^{n-1}e^{-\lambda w}dw =$
$\int_0^t Pr(t + s - w < X))\frac{\lambda^n}{(n-1)!}w^{n-1}e^{-\lambda w}dw =$

Inter-arrivals, X, are exponential. So we substitute the survival function.

$\int_0^t e^{-(\lambda(t+s-w))}\frac{\lambda^n}{(n-1)!}w^{n-1}e^{-\lambda w}dw =$
$\int_0^t e^{-(\lambda(t+s))}\frac{\lambda^n}{(n-1)!}w^{n-1}dw =$
$\{\frac{w^n}{n}\}_0^t \cdot e^{-(\lambda(t+s))}\frac{\lambda^n}{(n-1)!} =$
$\frac{t^n}{n} \cdot e^{-(\lambda(t+s))}\frac{\lambda^n}{(n-1)!} =$

$$\boxed{\frac{(\lambda t)^n \cdot e^{(\lambda t)}}{n!} \cdot e^{-s\lambda} = Pr(N(t) = n) \cdot Pr(W_{N(t)+1} > t + s)}$$

The second part holds true because we know that $W_{N(t)+1} > t$ thus we are looking only for the last interval time to be larger than s.

We have thus shown independence by proving that:
$Pr\{(N(t) = n)$ and $W_{N(t)+1} > t + s\} = Pr(N(t) = n) \cdot Pr(W_{N(t)+1} > t + s)$

6. Pinsky and Karlin, Problem 7.3.1

In another form of sum quota sampling (see Chapter 5, Section 5.4.2), a sequence of non negative independent and identically distributed random variables $X_1, X_2, \dots$ is observed, the sampling continuing until the first time that the sum of the observations exceeds the quota t. In renewal process terminology, the sample size is $N(t)+1$. The sample mean is:

$\frac{W_{N(t)+1}}{N(t)+1} = \frac{X_1+\dots+X_{N(t)+1}}{N(t)+1}$

An important question in statistical theory is whether or not this sample mean is unbiased. That is, how does the expected value of this sample mean relate to the expected value of, say, X_1? Assume that the individual X summands are exponentially distributed with parameter λ, so that $N(t)$ is a Poisson process, and evaluate the expected value of the foregoing sample mean and show that

$\mathbb{E}\{\frac{W_{N(t)+1}}{N(t)+1}\} = \frac{1}{\lambda}(1 - exp\{-\lambda t\})(1 + \frac{1}{\lambda t})$

SOLUTION:

From EXERCISE 7.3.3, the numerator and denominator are independent.

$\mathbb{E}\{\frac{W_{N(t)+1}}{N(t)+1}\} = \mathbb{E}(W_{N(t)+1}) * \mathbb{E}(\frac{1}{(N(t)+1)})$

By 7.7 in the book: $\mathbb{E}(W_{N(t)+1}) = \mu(1 + M(t))$

(By 7.7 and 7.6 and that $X_1 \sim Exponential(\lambda)$ and $W_n \sim Gamma(n, \lambda)$)
So $\mathbb{E}\{\frac{W_{N(t)+1}}{N(t)+1}\} = (\frac{1}{\lambda})(1 + (\lambda t)) * \mathbb{E}(\frac{1}{N(t)+1})$

$\mathbb{E}(\frac{1}{N(t)+1}) = \sum_{n=0}^{\infty}(\frac{1}{n+1})Pr(N(t) = n) = \sum_{n=0}^{\infty}(\frac{1}{n+1})\frac{(\lambda t)^n e^{(\lambda t)}}{n!} = \frac{1}{(\lambda t)}\sum_{n=0}^{\infty}(\frac{1}{n+1})(\lambda t)\frac{(\lambda t)^n e^{(\lambda t)}}{n!}$

Here, we are summing from 1 to infinity, so we need to subtract the probability at zero, giving us $(1 - exp(\lambda t))(\frac{1}{(\lambda t)})$ for the second term.

We thus have:
$\mathbb{E}\{\frac{W_{N(t)+1}}{N(t)+1}\} = (\frac{1}{\lambda})(1 + (\lambda t)) * \mathbb{E}(\frac{1}{N(t)+1}) = (\frac{1}{\lambda})(1 + (\lambda t))(1 - exp(\lambda t))(\frac{1}{(\lambda t)}) =$
$\boxed{\frac{1}{\lambda}(1 - exp\{-\lambda t\})(1 + \frac{1}{\lambda t})}$

7. Pinsky and Karlin, Problem 7.3.3

Pulses arrive at a counter according to a Poisson process of rate λ. All physically realizable counters are imperfect, incapable of detecting all signals that enter their detection chambers. After a particle or signal arrives, a counter must recuperate, or renew itself, in preparation for the next arrival. Signals arriving during the readjustment period, called dead time or locked time, are lost. We must distinguish between the arriving particles and the recorded particles.
The experimenter observes only the particles recorded; from this observation he desires to infer the properties of the arrival process. Suppose that each arriving pulse locks the counter for a fixed time τ. Determine the probability $p(t)$ that the counter is free at time t.

SOLUTION:
$Pr(\text{Free at time } t) = Pr(0 \text{ pulses between } t-\tau \text{ and } t)$

This is slightly more complicated, because we cannot consider negative time.
We want: $min(t-(t-\tau),0)$.

$$t-(t-\tau)_+ = \begin{cases} \tau & \text{if } t > \tau \\ t & \text{if } t < \tau \end{cases}$$

$$Pr(Poisson(\lambda \cdot min(t,\tau)) = 0) = \frac{(\lambda \cdot min(t,\tau))^0 e^{-(\lambda \cdot min(t,\tau))}}{0!} = \boxed{e^{-(\lambda \cdot min(t,\tau))}}$$

8. Pinsky and Karlin, Problem 7.3.4

This problem is designed to aid in the understanding of length-biased sampling. Let X be a uniformly distributed random variable on $[0, 1]$. Then, X divides $[0, 1]$ into the sub-intervals $[0, X]$ and $(X, 1]$. By symmetry, each sub-interval has mean length $\frac{1}{2}$.

Now pick one of these sub-intervals at random in the following way: Let Y be independent of X and uniformly distributed on $[0, 1]$, and pick the subinterval $[0, X]$ or $(X, 1]$ that Y falls in. Let L be the length of the subinterval so chosen. Formally:

$$\text{L} = \begin{cases} X & \text{if } Y \leq X \\ 1 - X & \text{if } Y > X \end{cases}$$

Determine the mean of L.

SOLUTION:

$E(L) = \int_0^1 Pr(Y < X)(x)dx + \int_0^1 Pr(Y > X)(1 - x)dx$

Because X and Y are independent, immediately integrate Y over X. Because y equals x when integrating from 0 to x and y equals (1-x) when integrating from x to 1, this looks like just regular substitution. The 'integrating' also has no effect on the length, as the lengths are only functions of X.

$E(L) = \int_0^1 Pr(Y < x)(x)dx + \int_0^1 Pr(Y > x)(1 - x)dx =$

$E(L) = \int_0^1 (x)^2 dx + \int_0^1 (1 - x)^2 dx =$

$E(L) = [\frac{(x)^3}{3}]_0^1 + [x - x^2 + \frac{(x)^3}{3}]_0^1 =$

$E(L) = \frac{(1)^3}{3} + \frac{(1)^3}{3} =$

$$\boxed{E(L) = \frac{2}{3}}$$

9. Pinsky and Karlin, Problem 7.3.5 (Recall that the Poisson process has $N(0) = 0$, so there is no bird at the far left.)

   Birds are perched along a wire as shown as below, according to a $PP(\lambda)$ At a fixed point t along the wire, let $D(t)$ be a random distance to the nearest bird. What is the mean of $D(t)$? What is the pdf $f_t(x)$ for $D(t)$?

   SOLUTION:

   We have guarantee that we have empty space to the left and right of the given time. If the side to the left intersects zero, then we need to cut it off so that it doesn't go negative.

$$Pr(D(t) > x) = Pr(\text{No birds in } (t-x, t+x]) = P(N(t+x) - N(t-x) = 0)$$

$$Pr(D(t) > x) = \begin{cases} P(N(t+x) = 0) = e^{-\lambda(x+t)} & \text{if } x > t \\ P(N(2x) = 0) = e^{-\lambda(2x)} & \text{if } x < t \end{cases}$$

   $E(D(t)) = \int_0^\infty Pr(D(t) > x)\, dx = \int_0^t e^{-2\lambda x}\, dx + \int_t^\infty e^{-\lambda(t+x)}\, dx =$

   Evaluating the first integral:

   $\int_0^t e^{-2\lambda x}\, dx = \left[-\frac{1}{2\lambda}e^{-2\lambda x}\right]_0^t = -\frac{1}{2\lambda}e^{-2\lambda t} + \frac{1}{2\lambda}$ For the second integral, let $z = t + x$ and $dz = dx$. When $x = t$, $z = 2t$, and when $x \to \infty$, $z \to \infty$.

   Thus:

   $\int_t^\infty e^{-\lambda(t+x)}\, dx = \int_{2t}^\infty e^{-\lambda z}\, dz = \left[-\frac{1}{\lambda}e^{-\lambda z}\right]_{2t}^\infty = \frac{1}{\lambda}e^{-2\lambda t}$

   Combining the results:

$$\boxed{E(D(t)) = \frac{1}{2\lambda}(1 + e^{-2\lambda t})}$$

   Taking the derivative of the CDF will give us the PDF:

   $\frac{d}{dx}(1 - e^{-\lambda(x+t)}) = \lambda e^{-\lambda(x+t)}$
   $\frac{d}{dx}(1 - e^{-2\lambda(x)}) = 2\lambda e^{-2\lambda(x)}$

$$\boxed{f_{D(t)}(x) = \begin{cases} \lambda e^{-\lambda(x+t)} & \text{if } x > t \\ 2\lambda e^{-\lambda(2x)} & \text{if } x < t \end{cases}}$$