Solutions Stochastic Processes and Simulation II, August 19, 2020

Problem 1: Poisson Processes

Consider a homogeneous Poisson process $\{N(t); t \geq 0\}$ with intensity $\lambda > 0$. Let $c > 0$ be a positive constant. Now color the point in the Poisson process independently red or blue where for $t > 0$ a point at position t is red with probability e^{-ct}, otherwise the point is blue.

a) What is the probability that the first point is red? (4pt)

Solution: The position of the first point has density $f(t) = \lambda e^{-\lambda t}$. So the probability that the first point is red is $\int_0^\infty f(t)e^{-ct}dt = \int_0^\infty \lambda e^{-(c+\lambda)t}dt = \lambda/(\lambda + c)$.

b) What is the distribution of the number of red points in $(0, \infty)$? (4pt)

Solution: Because the coloring of points is independent the red points themselves constitute a (non-homogeneous) Poisson Process with intensity λe^{-ct} (e.g. page 324 of Ross). So, the distribution of the number of points is Poisson distributed with expecation $\int_0^\infty \lambda e^{-ct}dtI\lambda/c$.

Denote the positions of the red points by $R_1 < R_2 < \cdots$, where $R_k = \infty$ if the total number of red points is less than k.

c) Let $s < t < \infty$. Compute $\mathbb{P}(R_1 < s|R_3 = t)$. (4pt)

Solution: Use the order statistic property for non-homogeneous Poisson processes (e.g. Ross page 675) And note that the number of points in $(0, t)$ is 2. The positions of those points are distributed as two independent random variables with density

$$g(s) = \frac{\lambda e^{-cs}}{m(t)} = \frac{\lambda e^{-cs}}{\int_0^t \lambda e^{-cu}du} = \frac{ce^{-cs}}{1 - e^{-ct}}.$$

So, for $s < t$, noting that $R_1 > s$ means that there are no points in $(0, s)$.

$$\begin{aligned}\mathbb{P}(R_1 < s|R_3 = t) = 1 - \mathbb{P}(R_1 > s|R_3 = t) = 1 - (1 - \int_0^s g(u)du)^2 \\ = 1 - \left(1 - \frac{\int_0^s ce^{-cu}du}{1 - e^{-ct}}\right)^2 = 1 - \left(1 - \frac{1 - e^{-cs}}{1 - e^{-ct}}\right)^2\end{aligned}$$

Problem 2: Renewal Theory

A quiz show candidate is asked a series of questions. Each answer can be evaluated to be either C (for correct) or F (for false). The sequence of answers can be described by a Markov chain. The first answer is correct with probability p_1. A correct answer is directly followed by another correct answer with probability p_C, while a false answer is directly followed by a correct answer with probability p_F. Let $C(n)$ be the number of correct answers among the first n answers.

a) What is the expected number of questions needed to obtain the first correct answer? (4pt)

Solution: Let Z be the positition of the first correct answer. Then $\mathbb{P}(Z = 1) = p_1$ and for $k > 1$ we have $\mathbb{P}(Z = k) = (1 - p_1)(1 - p_F)^{k-2}p_F$ So,

$$\begin{aligned}\mathbb{E}[Z] &= \sum_{k=1}^{\infty} k\mathbb{P}(X = k) = p_1 + \frac{1 - p_1}{1 - p_F} \sum_{k=2}^{\infty} k(1 - p_F)^{k-1} p_F \\ &= p_1 + \frac{1 - p_1}{1 - p_F} \sum_{k=1}^{\infty} k(1 - p_F)^{k-1} p_F - \frac{1 - p_1}{1 - p_F} p_F = p_1 + \frac{1 - p_1}{1 - p_F} \frac{1}{p_F} - \frac{1 - p_1}{1 - p_F} p_F \\ &= \frac{p_F(p_1 - p_F) + (1 - p_1)}{(1 - p_F)p_F} = \frac{1 - p_1 + p_F}{P_F}.\end{aligned}$$

b) What is the (almost sure) long-run fraction of answers that is correct? That is, what is the almost sure limit of $C(n)/n$? (4pt)

Hint: Note that p_1, p_C and p_F are not necessarily equal.

Solution: Life would be easier if first answer would be correct with probability p_C, because then we can say that there is a renewal every time a correct answer is given. However, for the long run the probability that the first answer is correct does not matter. So let us assume that $p_1 = p_C$ and the expected number of questions after a renewal until a new renewal is $\frac{1-p_C+p_F}{p_F}$, and the (almost sure) long-run fraction of answers that is correct is one divided by this number: $\frac{p_F}{1-p_C+p_F}$.

c) Let $X(n)$ be the number of subsequent correct answers given at time n. So, if the sequence starts

$$C, C, F, C, F, \cdots,$$

then

$$X(1) = 1, \quad X(2) = 2, \quad X(3) = 0, \quad X(4) = 1 \quad \text{and} \quad X(5) = 0.$$

For all $n \in \{1, 2, \cdots\}$, assume that at time n the candidate receives a reward of $X(n)$. What is the expected long run income of the candidate per time unit? That is what is

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} X(j) \quad ?$$

(4pt)

Solution: Use Renewal reward theory, but now with a renewal every time an incorrect answer is given and the answer of an incorrect answer at the first question is $1 - p_F$. Let Y be the position of the first incorrect answer. As for Z we obtain $\mathbb{P}(Y = 1) = 1 - p_F$ and for $k > 1$ we have $\mathbb{P}(Y = k) = p_F p_C^{k-2}(1 - p_C)$. Then according to renewal reward theory we are interested in

$$\frac{1}{\mathbb{E}[Y]} \mathbb{E}[\sum_{j=1}^{Y-1} X(j)] = \frac{\mathbb{E}[Y(Y-1)]}{2\mathbb{E}[Y]}$$

We otbain by changing the role of p_F and $1 - p_C$.

$$\mathbb{E}[Y] = \frac{1 - p_C + p_F}{1 - p_C}$$

and

$$\mathbb{E}[Y(Y-1)] = \frac{p_F}{2(1 - p_C)^2}.$$

Problem 3: Queueing Theory

Consider the following $M/M/\infty$ queue with catastrophes: Customers arrive in the system according to a homogeneous Poisson Process with intensity $\lambda > 0$. Upon entering the system, a customer immediately receives service from one of the infinitely many servers. Workloads are independent and exponentially distributed with rate $\mu > 0$ (So, with expectation $1/\mu$). A customer who has been in the system for the duration of his or her workload, immediately leaves the system. In addition catastrophes occur according to a Poisson process with intensity δ. This Poisson process is independent of the arrival process of customers. At the moment of a catastrophe, all customers present leave the system immediately. Workloads are independent of both the arrival process and the times of catastrophes.

a) Provide the expected time a customer is in the system? (2pt)

Solution: Say that a customer has workload Y, which is exponentially distributed with mean $1/\mu$ and that the time until next catastrophe from the moment of his or her arrival is X, which is independent of Y and exponentially distributed with mean $1/\delta$. So the time until leaving is $\min(X, Y)$ which is exponentially distributed with parameter $\delta + \mu$ and thus with expectation $W = 1/(\delta + \mu)$

b) What is the long run average number of customers in the system? (3pt)

Solution: Using $L = \lambda W$ (Page 483 of Ross) and thus $L = \lambda/(\delta + \mu)$.

c) For $k \in \{1, 2, \cdots\}$ Let N_k be the number of customers that leaves the system at the k-th catastrophe. What is $\lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} N_k$? (3pt)

Solution: Since catastrophes come as Poisson arrivals, we can use PASTA and the answer is L.

d) Let $\mathcal{S} = \{0, 1, 2, \cdots\}$ be the "state space" of the process describing the number of customers in the system. Let $\{P_s\}_{s\in\mathcal{S}}$ be the stationary distribution of the Queueing system described in this problem. Provide the "balance equations" characterizing this stationary distribution. That is, provide relations beween the P_s ($s \in \mathcal{S}$), which are in theory enough to compute them all. (4pt)

Hint: Note you do not have to solve the balance equations.

Solution: The rate into state 0 is: $\mu P_1 + \delta(1 - P_0)$ and the rate out is λP_0. In general for $k \geq 1$ the rate into state k is $\lambda P_{k-1} + \mu(k+1)P_{k+1}$ while the rate out is $(\delta + \lambda + k\mu)P_k$. So the balance equations are:

$$\mu P_1 + \delta(1 - P_0) = \lambda P_0$$

and for $k \geq 1$

$$\lambda P_{k-1} + \mu(k+1)P_{k+1} = (\delta + \lambda + k\mu)P_k.$$

Problem 4: Brownian Motion and Stationary Processes

Let $\{B(t); t \geq 0\}$ be a standard Brownian motion (So we assume $B(0) = 0$ and $Var(B(1)) = 1$). Let

$$m(t) = \min_{0 \leq s \leq t} B(t).$$

a) Argue that for $x, y > 0$ we have

$$\mathbb{P}(m(t) < -x, B(t) > y - x) = \mathbb{P}(B(t) > y + x).$$

(4pt)

Solution: Use reflection principle in $-x$. So at the first hitting time of $-x$ (say T_{-x}) the probability of going from $-x$ to above $y-x$ (i.e. an increase of at least y) in $t - T_{-x}$ time units is equal to going from $-x$ to below $-(y+x)$ (i.e. a decreas of at least y) in that time interval. Then further use that the Brownian motion is symmetric so $\mathbb{P}(B(t) > y + x) = \mathbb{P}(B(t) < -(y + x))$.

b) Again for $x, y > 0$, compute $\mathbb{P}(B(t) > y - x | m(t) > -x)$. (4pt)

Hint: You may use a) to obtain $\mathbb{P}(m(t) > -x, B(t) > y - x)$.

Solution:

$$\begin{aligned}
\mathbb{P}(B(t) > y - x | m(t) > -x) &= \frac{\mathbb{P}(B(t) > y - x, m(t) > -x)}{\mathbb{P}(m(t) > -x)} \\
&= \frac{\mathbb{P}((B(t) > y - x) - \mathbb{P}(B(t) > y - x, m(t) < -x)}{\mathbb{P}(m(t) > -x)} \\
&= \frac{\mathbb{P}(B(t) > y - x) - \mathbb{P}(B(t) > y + x)}{1 - 2\mathbb{P}(B(t) > x)} = \frac{\int_{y-x}^{y+x} \frac{1}{\sqrt{2\pi t}} e^{-s^2/(2t)} ds}{2 \int_0^x \frac{1}{\sqrt{2\pi t}} e^{-s^2/(2t)} ds}
\end{aligned}$$

c) Show that

$$\lim_{\epsilon \searrow 0} \mathbb{P}(B(t) > y | m(t) > -\epsilon) = e^{-y^2/(2t)}$$

and compute

$$\lim_{\epsilon \searrow 0} \mathbb{E}[B(t) | m(t) > -\epsilon].$$

Here $\lim_{\epsilon \searrow 0}$ means the limit as ϵ decreases to 0. (4pt)

Solution: Using part b with $x = \epsilon$ we obtain

$$\mathbb{P}(B(t) > y | m(t) > -\epsilon) = \frac{\int_y^{y+2\epsilon} \frac{1}{\sqrt{2\pi t}} e^{-s^2/(2t)} ds}{2 \int_0^\epsilon \frac{1}{\sqrt{2\pi t}} e^{-s^2/(2t)} ds} = \frac{2\epsilon \frac{1}{\sqrt{2\pi t}} e^{-y^2/(2t)} + o(\epsilon)}{2\epsilon \frac{1}{\sqrt{2\pi t}} + o(\epsilon)} \to e^{-y^2/(2t)}.$$

and because $\int_0^\infty \frac{1}{\sqrt{2\pi t}} e^{-y^2/(2t)} dy = 1/2$, we have

$$\lim_{\epsilon \searrow 0} \mathbb{E}[B(t) | m(t) > -\epsilon] = \int_0^\infty \mathbb{P}(B(t) > y | m(t) > -\epsilon) dy = \int_0^\infty e^{-y^2/(2t)} dy = \frac{1}{2}\sqrt{2\pi t}.$$

Problem 5: Simulation

Let X and Y be random variables with density functions

$$f_X(x) = \begin{cases} \sqrt{\frac{2}{\pi}} e^{-x^2/2} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \quad \text{and} \quad f_Y(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}.$$

Assume that it is easy to simulate a realisation of the random variable Y.

a) If you use the rejection method using $f_Y(x)$ to simulate a random variable with density function $f_X(x)$, which value of $\lambda \in (0, \infty)$ gives the lowest possible probability of rejection of a proposed realisation of X? (6pt)

Remark: The rejection method above can be used (after an extra step) to simulate a realisation of a standard normal random variable.

Solution: Using the theory on the rejection method we first need to find

$$C_\lambda = \max_{x>0} \frac{f_X(x)}{f_Y(x)} = \max_{x>0} \frac{\sqrt{\frac{2}{\pi}} e^{-x^2/2} e^{\lambda x}}{\lambda} = \max_{x>0} \sqrt{\frac{2}{\pi \lambda^2}} e^{\lambda^2/2} e^{-(x-\lambda)^2/2} = \sqrt{\frac{2}{\pi \lambda^2}} e^{\lambda^2/2}.$$

Since the rejection probability is C_λ we are interested in the value of λ for which C_λ is minimal. That is the value of λ for which $e^{\lambda^2/2}/\lambda$ is minimal, which in turn is minimal (by taking logarithms when $g(\lambda) = \lambda^2/2 - \log \lambda$ is minimal. Now $g'(\lambda) = \lambda - 1/\lambda$ which is 0 only if $\lambda = 1$, which is indeed where C_λ takes its minimum. So for $\lambda = 1$, C_λ is minimal and the rejection method gives the lowest possible probability of rejection.

Let $\mathbb{N}$ be the set of strictly positive integers and For every $k \in \mathbb{N}$, define $[k] = \mathbb{N} \cup (0, k]$ as the set of positive integers not exceeding k.

Consider the following construction of the random process $\{W(t); t \in [0, 1]\}$.
Only Steps 1, 2 and equation (1) **are needed for solving part b). You can safely ignore the other steps!**

1. For all $n \in \mathbb{N}$ and $j \in [2^{n-1}]$, let N_0 and $N_{n,j}$ be independent standard normal distributed random variables.
2. Set $W(0) = 0$ and $W(1) = N_0$.
3. Assume that you know $W(i2^{-(n-1)})$, for all $i \in [2^{n-1}]$ (which you do for $n = 1$). Then for all $j \in [2^{n-1}]$ define
$$W((2j-1)2^{-n}) = \frac{W((j-1)2^{-(n-1)}) + W(j2^{-(n-1)})}{2} + 2^{-(n+1)/2} N_{n,j}.$$
In this way you can find $W(x)$ for all $x \in [0, 1]$ with finite binary representation. In particular, $n = j = 1$ gives
$$W(1/2) = \frac{W(0) + W(1)}{2} + 2^{-1} N_{1,1}. \qquad (1)$$
4. For $n \in \mathbb{N}$, Define $\{W_n(t); t \in [0, 1]\}$ by connecting for $i \in [2^n]$, the points $((i-1)2^{-n}, W((i-1)2^{-n}))$ and $(i2^{-n}, W(i2^{-n}))$ by straight line segments. That is, for $i \in [2^n]$ and $t \in [0, 2^{-n}]$ define
$$W_n((i-1)2^{-n} + t) = W((i-1)2^{-n}) + 2^n t[W(i2^{-n}) - W((i-1)2^{-n})].$$
5. $\{W(t); t \in [0, 1]\}$ is the pointwise limit of $\{W_n(t); t \in [0, 1]\}$ as $n \to \infty$.

b) Deduce from the definition of $W(0), W(1)$ and $W(1/2)$ that $W(1/2)$ has a normal distribution with expectation 0 and variance 1/2. Show further that $W(1/2)$ and $W(1) - W(1/2)$ are independent. (6pt)

Solution: Note
$$W(1/2) = \frac{W(1) + N_{1,1}}{2} = \frac{N_0 + N_{1,1}}{2} \quad \text{and} \quad W(1) - W(1/2) = W(1) - \frac{W(1) + N_{1,1}}{2} = \frac{N_0 - N_{1,1}}{2}.$$
Since N_0 and $N_{1,1}$ are normal distributed and independent, and because the sum of two independent random variables is normal as well, with expectation the sum of the expectations of the summands and with variance the sum of the expectations of the summands. $W(1/2)$ and $W(1)-W(1/2)$ are both normal distributed with expectation $(0+0)/2$ and variance $(1+1) \times 1/4$.

Two normal distributed random variables are independent if they have 0 covariance. That $W(1/2)$ and $W(1) - W(1/2)$ have 0 covariance follows from

$$\begin{aligned} Cov\left(\frac{N_0 + N_{1,1}}{2}, \frac{N_0 - N_{1,1}}{2}\right) &= \frac{1}{4} Cov(N_0 + N_{1,1}, N_0 - N_{1,1}) \\ &= \frac{1}{4}\left(Cov(N_0, N_0) - Cov(N_{1,1}, N_{1,1}) - Cov(N_0, N_{1,1}) + Cov(N_0, N_{1,1})\right) \\ &= Var(N_0) - Var(N_{1,1}) = 0 \end{aligned}$$