Name: Key

STAT 521 – Test 2 – Spring 2019

1. Assume a population of amoebas reproduces according to a branching process as follows. Amoebas reproduce independently, and each amoeba has j offspring with probability P_j for $j = 0, 2$, where $P_0 = 0.35$ and $P_2 = 0.65$.

(a) Find the expected number of offspring for a random amoeba. Show work.

$$\mu = (0)(0.35) + (2)(0.65) = 1.30$$

(b) If the initial population size is 1, then find the expected population size at the 7th generation.

$$E(X_7) = (1.30)^7 = 6.275$$

(c) Find the probability that the population eventually dies out. Show your work.

$$\pi_0 = 0.35 + 0.65\,\pi_0^2$$

$$0 = 0.65\,\pi_0^2 - \pi_0 + 0.35$$

$$\pi_0 = \frac{1 \pm \sqrt{1 - 4(.65)(.35)}}{1.30}$$

$\Rightarrow \pi_0$ is 1 or 0.538

$$\Rightarrow \boxed{\pi_0 = 0.538}$$

(d) If we are told that the zero-th generation has 1 amoeba AND we are told that single initial amoeba has 2 offspring, what is the expected population size at the 3rd generation? Briefly explain.

$X_0 = 1$, $X_1 = 2$. Consider these two branches each going 2 further generations into the future to reach the 3rd generation.

Each branch has expected size $(1.30)^2 = 1.69$

So $E(X_3) = 1.69 + 1.69 = 3.38$

2. Consider a continuous random variable X for which you know that $P(X > t + s \mid X > s) = P(X > t)$ for all $s, t \geq 0$.

(a) We say that such a random variable has the memoryless property.

(b) Suppose that this random variable has $E(X) = 10$. Find $P(X > 5.5 \mid X > 1)$. Show work.

$$P(X>5.5 \mid X>1) = P(X>4.5) = e^{-4.5/10}$$

$$= e^{-0.45} = \boxed{.6376}$$

3. Suppose we want to use the Metropolis-Hastings algorithm to sample from a Lomax distribution with pdf

$$f(x) = 3(1+x)^{-4}, \; x > 0$$

(and 0 otherwise).

(a) Briefly explain why the exponential distribution, with mean equal to the current value in the chain, would be a reasonable choice of proposal distribution.

It is a continuous distribution which has support on $(0, \infty)$, the same support as the target distribution.

(b) If we use the proposal distribution suggested in (a), give an expression for the probability of accepting the next proposed (candidate) value in the chain. Simplify it, where possible.

$$q(i,j) = \frac{1}{i} e^{-j/i} \qquad q(j,i) = \frac{1}{j} e^{-i/j}$$

$$\Rightarrow \frac{\pi(j)\, q(j,i)}{\pi(i)\, q(i,j)} = \frac{3(1+j)^{-4} \frac{1}{j} e^{-i/j}}{3(1+i)^{-4} \frac{1}{i} e^{-j/i}} = \frac{i(1+i)^4 e^{-i/j}}{j(1+j)^4 e^{-j/i}}$$

$$\Rightarrow \alpha(i,j) = \min\left\{ \frac{i(1+i)^4 e^{-i/j}}{j(1+j)^4 e^{-j/i}}, 1 \right\}$$

(c) Why would the Gibbs sampler not be an appropriate method to sample from the distribution in this problem?

The target distribution is not multivariate.

4. Suppose that at three restaurants at the student union, Alan is being served at Arby's, Bob is being served at Burger King, and Cindy is being served at Chick-Fil-A. Suppose that the waiting times at the three restaurants are: exponential with **mean** 4 minutes for Arby's; exponential with **mean** 2 minutes for Burger King; and exponential with **mean** 8 minutes for Chick-Fil-A. And suppose all the waiting times are independent.

(a) What is the probability that Alan is the first of the three to finish being served? Show work.

$$P\left[X_A = \min_j X_j\right] = \frac{\frac{1}{4}}{\frac{1}{4}+\frac{1}{2}+\frac{1}{8}} = \boxed{.2857}$$

(b) Would the probability in part (a) change if you were told that Alan started being served 5 minutes ago, while Bob and Cindy started being served only 20 seconds ago? **Briefly** explain why or why not.

No, because exponential r.v.'s have the memoryless property.

(c) What is the expected time until the first customer of the three (whoever is first to finish) is done being served?

$$\min_i X_i \sim \text{expon}\left(\tfrac{1}{4}+\tfrac{1}{2}+\tfrac{1}{8}\right)$$

$$\Rightarrow \min_i X_i \sim \text{expon}\left(\tfrac{7}{8}\right)$$

$$\Rightarrow E(\min X_i) = \frac{8}{7} = 1.14 \text{ minutes.}$$

(d) Now suppose that Alan does indeed finish first, and the other two customers, Bob and Cindy, are still waiting to finish. Suppose then that one of these two finishes being served. At that time, find the expected additional waiting time of the final remaining customer (i.e., the expected time between when the second customer leaves and when the final customer leaves). Show work.

Let W_R = waiting time of remaining customer

$$E(W_R) = E(W_R \mid \text{Bob done before Cindy})\,P(\text{Bob done before Cindy}) + E(W_R \mid \text{Cindy done before Bob})\,P(\text{Cindy done before Bob})$$

$$= (8)\left(\frac{\frac{1}{2}}{\frac{1}{2}+\frac{1}{8}}\right) + (2)\left(\frac{\frac{1}{8}}{\frac{1}{2}+\frac{1}{8}}\right)$$

$$= (8)(0.8) + (2)(0.2) = \boxed{6.8 \text{ minutes}}$$

(e) Consider a similar situation, except that it is Sunday and Chick-fil-A is closed. Alan is being served by the one cashier at Arby's, and Bob is being served by the one cashier at Burger King. Cindy enters the student union and decides to wait and go to whichever restaurant (Arby's or Burger King) has its cashier become available first. (Only Alan and Bob are being served, no other customers.) Find the probability that Cindy takes longer than 3 minutes to be served, **not including the time** she spends waiting for a cashier to become available.

Let S_c = Cindy's service time

$$P(S_c > 3) = P(S_c > 3 \mid \text{Alan done first})\,P(\text{Alan done first}) + P(S_c > 3 \mid \text{Bob done first})\,P(\text{Bob done first})$$

$$= \left(e^{-3/4}\right)\left(\frac{\frac{1}{4}}{\frac{1}{4}+\frac{1}{2}}\right) + \left(e^{-3/2}\right)\left(\frac{\frac{1}{2}}{\frac{1}{4}+\frac{1}{2}}\right)$$

$$(.4724)(.3333) + (.2231)(.6667) = \boxed{.3062}$$

5. Suppose that the claims received by a large insurance company follow a Poisson process with rate λ = 2.5 per hour.

(a) The company begins tracking received claims at 8:00 a.m. What is the probability of exactly 2 claims arriving between 9:00 and 9:45 a.m.? Show work.

$$P[N(1.75) - N(1) = 2] = P[N(0.75) = 2]$$

$$= \frac{e^{-(2.5)(0.75)}\,[(2.5)(0.75)]^2}{2!} = \boxed{0.2696}$$

(b) Suppose the first claim of the day arrived at 8:15 a.m. What is the probability that the second claim of the day arrives before 8:45 a.m.? Show work.

$T_2 \sim \text{expon}(2.5)$

$$P[T_2 < 0.5 \text{ hours}]$$

$$= 1 - e^{-(2.5)(0.5)} = \boxed{0.7135}$$

(c) Find the expected arrival time (i.e., in terms of the time on the clock) of the 18th claim of the day.

$$S_{18} \sim \text{gamma}(18, 2.5) \Rightarrow E(S_{18}) = \frac{18}{2.5} = 7.2$$

$$\Rightarrow \boxed{3{:}12 \text{ pm}}$$

(d) Suppose that the second claim received in the afternoon (i.e., the second claim arrival past noon) arrived at 12:40 p.m. What is the probability that the first arrival of the afternoon happened between 12:25 and 12:30 p.m.?

First arrival $\sim$ Unif $(0, 40)$ minutes past noon

$$\frac{30-25}{40-0} = \frac{5}{40} = \boxed{.125}$$

(e) Suppose that 20 percent of the claims involve car accidents and 80 percent do not (and that whether different claims involve car accidents or not is independent across claims). Find the probability that exactly 3 claims involving car accidents are received in the first 6 hours of a day. Show work.

$$N_1(t) \sim \text{Poisson}\left((0.20)(2.5)\right) \Rightarrow \text{Pois}(0.5)$$

$$P[N_1(6) = 3] = \frac{[(0.5)(6)]^3 e^{-(0.5)(6)}}{3!}$$

$$= \boxed{0.224}$$

Extra credit [2 points]: Consider the sum of a RANDOM number of independent exponential r.v.'s. The distribution of this sum is named for which statistician (and from which country is he?).

David Cox (Great Britain)