Stat 134: Review Session 6.1-6.3

Adam Lucas

December 5th, 2018

Conceptual Review

Let $-\infty \leq a < b \leq \infty$.

a. Suppose X is a continuous random variable with possible values on (a, b). Let A be an event. Write down a formula for $P(A)$ by conditioning on X.

b. Let Y be a random variable, and let $B_1, B_2, B_3, \ldots$ be events. Write down an expression for $E(Y)$ by conditioning on the events B_i.

c. Let X_1, X_2 be random variables, and suppose h is a two-input function, and consider $\theta = E[h(X_1, X_2)]$. Now let g be a function defined by $g(x) = E[h(x, X_2)] - \theta$. Explain why $g(X_1)$ is not 0 but is a random variable with $E[g(X_1)] = 0$.

d. Let $N \sim Geometric(p)$ on $\{1, 2, \ldots\}$. What is $E[cN + d | N > m]$?

Problem 1:

You and your friend are going to play a rock-paper-scissors game all day. You know that you choose rock 1/3 of the time, paper 1/6 of the time, and scissors 1/2 of the time. However, your friend does not like paper. Instead, your friend plays rock 1/2 of the time and scissors half of the time. Each time you duel, each player randomly picks one of rock, paper, or scissors independently. A duel corresponds to a pair, for example (Rock, Rock).

a. Find the expected number of duels it takes you to get:
   i. (Rock, Rock)
   ii. (Scissor, Scissor)
   Think of each duel as a trial.
   $P((\text{Rock}, \text{Rock})) = P(\text{you choose rock}) \cdot P(\text{friend chooses rock}) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$.
   Then the expected number of duels to get (R, R) is 6.
   $P((\text{Scissor}, \text{Scissor})) = P(\text{you choose scissor}) \cdot P(\text{friend chooses scissor}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.
   Then the expected number of duels to get (S, S) is 4.

b. Find the expected number of duels it takes to get either two (Rock, Rock) or two (Scissor, Scissor) in a row, whichever comes first. After a (R,R) stalemate occurs, neither player will pick Scissor. Instead, Paper absorbs this probability. The same works for a (S, S) stalemate.

Let W be the number of duels required. Condition on the outcome of the first duel to get

$$E(W) = \frac{1}{6}E[W|(R,R)_1] + \frac{1}{4}E[W|(S,S)_1] + (1 - \frac{1}{6} - \frac{1}{4})E[W| \text{ not } (R,R)_1 \text{ or } (S,S)_1] =$$

$$= \frac{1}{6}\left[\frac{1}{6}E[W|(R,R)_1,\ (R,R)_2] + \frac{5}{6}E[W|(R,R)_1,\ \text{not } (R,R)_2]\right] +$$

$$+ \frac{1}{4}\left[\frac{1}{4}E[W|(S,S)_1,\ (S,S)_2] + \frac{3}{4}E[W|(S,S)_1,\ \text{not } (S,S)_2]\right] + \frac{7}{12}[1 + E(W)].$$

Therefore

$$E(W) = \frac{1}{6}\left[\frac{1}{6}(2) + \frac{5}{6}(2 + E(W))\right] + \frac{1}{4}\left[\frac{1}{4}(2) + \frac{3}{4}(2 + E(W))\right] + \frac{7}{12}[1 + E(W)].$$

By performing the calculations, we get $E(W) = \frac{17}{12} + \frac{131}{144}E(W)$, so $E(W) = \frac{204}{13} \approx 15.7$.

Problem 2: Conjugate Priors "Gamma-Poisson"

Suppose the number of car accidents on the Bay Bridge every year, X, follows a $Poisson(\mu)$ distribution given μ. The rate μ is a realization from another distribution (called a prior) $M \sim Gamma(r, \lambda)$.

a. Given that X = k, what is the posterior distribution of the rate μ? (i.e. find the distribution of $M|X = k$)

Note that m is the variable here, and k is a constant. We do not need to explicitly find $P(X = k)$.

$$f_{M|X}(m) * dm = P(M \in dm|X = k) = \frac{P(M \in dm, X = k)}{P(X = k)} = \frac{P(X = k|M \in dm) * P(M \in dm)}{P(X = k)}$$

$$= \frac{\frac{e^{-m}m^k}{k!} * \frac{\lambda^r}{\Gamma(r)} m^{r-1} e^{-\lambda m}}{P(X = k)} * dm$$

$$\therefore f_{M|X}(m) = e^{-m-\lambda m} * m^{k+r-1} * \frac{\lambda^r}{k!\Gamma(r)P(X = k)} \quad \propto_m \quad m^{(k+r)-1}e^{-(\lambda+1)m} \sim Gamma(k + r, \lambda + 1)$$

This means that once we observe one data point (X = number of car accidents this year), we have gained information about μ.

b. Compare the expectation of M and $M|X = k$.

Plug in $r = 1,\ \lambda = 1,\ k = 5$

$$E(M) = r * \frac{1}{\lambda} = 1 \qquad E(M|X = k) = (r + k) * \frac{1}{\lambda + 1} = \frac{6}{2} = 3 > 1$$

Our previous thought about μ was a distribution around 1. After seeing that we got 5 accidents this year, our perspective about μ has moved towards the right, centering around 3 now. This represents a Bayesian inference model for μ.

Problem 3: Branching Process

Suppose there is a certain species that reproduces as follows: At time $n = 0$, there is one organism, and for each time step $n \geq 0$, each organism independently gives birth to 0 organisms with probability 1/4, or 2 organisms with probability 3/4, then dies. Let X_n denote the number of organisms alive at time n. We wish to find the probability that the species goes extinct (and thus the probability that it survives forever).

a. Show that $E(X_{n+1}) = \frac{3}{2}E(X_n)$. Note that this implies that $E(X_n) = (3/2)^n$.

   Let K denote the distribution given by $P(K = 0) = 1/4$, and $P(K = 2) = 3/4$.
   First, observe that $E(X_{n+1}|X_n = m) = m \cdot \frac{3}{2}$.
   This is because we are given that the n-th generation has m organisms, and each organism will give birth to K organisms, which clearly has expectation $E(K) = 3/2$.
   Therefore $E(X_{n+1}) = E[E(X_{n+1}|X_n)] = \frac{3}{2}E(X_n)$.

b. Let $u_n = P(X_n = 0)$. Show that for $n \geq 1$, $u_n = \frac{1}{4} + \frac{3}{4}u_{n-1}^2$.

   Perform a first-step analysis and use the law of total probability to get

   $$u_n = P(X_n = 0) = P(X_n = 0|K = 0)P(K = 0) + P(X_n = 0|K = 2)P(K = 2) = \frac{1}{4} + \frac{3}{4}u_{n-1}^2$$

   The second term above follows from the fact that the organisms are independent and thus start their own family tree. That is, if the first organism gives birth to two organisms, then we can consider each organism separately and notice that the process from here is exactly the same as the original one, except that we are now one step closer to time step n.

c. It may take some additional analysis to show that the sequence $\{u_n\}$ converges (indeed, it is an increasing sequence bounded above by 1). Taking this for granted, what is the probability that the species will eventually go extinct? *Hint:* You should arrive at 2 possible answers. Argue intuitively why one of them should be the correct answer.

   We assume that $u = \lim_{n\to\infty} u_n$ exists. Note that u_n is the probability that the species goes extinct by time n, so it follows that u is the probability that the species goes extinct eventually. From part (b), we must have $u = \frac{1}{4} + \frac{3}{4}u^2$ by taking $n \to \infty$. We can write

   $$\frac{3}{4}u^2 - u + \frac{1}{4} = 0 \implies 3u^2 - 4u + 1 = 0 \implies (3u - 1)(u - 1) = 0.$$

Thus there are two possible solutions: $u = \frac{1}{3}$ and $u = 1$. Since $E(X_n) = (3/2)^n \to \infty$ as $n \to \infty$, it should be likely that the species survives forever. Therefore, $u = 1/3$ and in fact the chance the species survives forever is $2/3$.