1. Let $S_n = 2 + \sum_{i=1}^n X_i$ with $X_i \in \{-1, 1\}$ and $P(X_i = -1) = \frac{1}{2} = P(X_i = 1)$.

   (a) What is ES_n?

   > **Solution:** $E2 + \sum_{i=1}^n EX_i = 2 + 0 = 2.$

   (b) Let X be a random variable and A be an event. The *conditional expectation* formula is

$$EX = E[X \mid A]P(A) + E[X \mid A^c]P(A). \tag{1}$$

   Let N be the smallest n such that $S_n \in \{1, k\}$. So N is the first time that $S_n = 1$ or $S_n = k$. Let $p = P(S_N = 1)$. Use (1) to write a formula for ES_N in terms of p.

   > **Solution:** $ES_N = p \cdot 1 + (1 - p)k.$

   (c) Using the fact (called the optional stopping theorem) that $ES_N = 0$. Solve for p.

   > **Solution:** $0 = p + (1 - p)k.$ so $p = \frac{1}{k}$.

2. Let $X = \text{geometric}(p)$.

   (a) What is $P(X = k)$?

   > **Solution:** $(1 - p)^{k-1}p$

   (b) What is $P(X = k + j \mid X > j)$?

   > **Solution:** We use the definition of $P(X = k + j \mid X \geq j)$:
   >
   > $$\frac{P(\{X = k + j\} \cap \{X > j\})}{P(X > j)} = \frac{P(X = k + j)}{(1 - p)^j} = \frac{(1 - p)^{k+j-1}p}{(1 - p)^j} = (1 - p)^{k-1}p.$$

   (c) The similarity between the previous two answers is referred to as the "memoryless property." Write in words why this is a fitting description.

   > **Solution:** This says that even if we know the wait time is at least j, the remaining time still has the same distribution. So there is no memory of the past time spent waiting.

   (d) Let $X = \text{geometric}(1/3)$ and $Y = \text{geometric}(p)$. Find p so that $P(X \leq Y) = \frac{2}{3}$. Explain where you use the memoryless property.

   > **Solution:** Set $q = P(X \leq Y)$. Let $A_1 = \{X = 1\}$, $A_2 = \{X > 1\}$. We can use the recursion from Example 3.17 to write
   >
   > $$q = 1 \cdot \frac{1}{3} + 0 \cdot \frac{2}{3}q + q\frac{2}{3}(1 - p).$$
   >
   > The memoryless property is in the last term. We are assuming both X and Y are greater than 1. Set $q = 2/3$ and we have
   >
   > $$\frac{2}{3} = \frac{1}{3} + \frac{2}{3}\frac{2}{3}(1 - p)$$
   >
   > This gives $p = 1/4$.

3. Let $X_i \in \{-1, 1\}$ with $P(X_i = 1) = p$ and $P(X_i = -1) = 1 - p$. Let $S_0 = 0$ and define $S_n = \sum_{i=1}^n X_i$. Let N be the random variable for the smallest number such that $S_N = -1$.

   (a) Let $A_i = \{S_1 = i\}$. For which i is $P(A_i) > 0$?

   Solution: $i = -1, i = 1$

   (b) What is $P(A_i)$ for i from the previous part?

   Solution: $P(A_{-1}) = 1 - p$, $P(A_1) = p$.

   (c) Condition on the events A_i to write an expression involving EN, p, and $E[N \mid A_1]$.

   Solution:

   $$EN = E[N \mid A_{-1}]P(A_{-1}) + E[N \mid A_1]P(A_1) = (1 - p) + E[N \mid A_1]p.$$

   (d) Explain in words why $E[N \mid A_1] = 2EN$.

   Solution: If $S_1 = 1$ then the process has to get back to 0, then get back to -1. The time to do both of these things is the same as adding two independent copies of N. Taking expectations gives the claimed formula.

   (e) Make the substitution from (d) and solve for EN.

   Solution: $EN = 1 - p + 2pEN$ hence

   $$EN = \frac{1 - p}{1 - 2p}.$$

   (f) What is EN when $p = \frac{1}{4}$.

   Solution: $\frac{3/4}{1/2} = 6/4$.

   (g) What is EN when $p = \frac{1}{2}$?

   Solution: ∞.

   (h) What is EN when $p > 1/2$? (*Hint: use the formula from (e) before isolating EN*)

   Solution: When $p > 1/2$ we get the nonsense answer EN is equal to something larger than it. The only way this can be true is if $EN = \infty$.

4. (This was taken from an actuarial exam.) An insurance company designates 10% of its customers as high risk and 90% as low risk. The number of claims made by high and low risk customers in a calendar year are Poisson distributed with means $\lambda_h = 2$ and $\lambda_\ell = 1$ and is independent of the number of claims made by a customer in the previous calendar year. Calculate the expected number of claims made in calendar year 2017 by a customer who made one claim in calendar year 2016.

Solution: Replace .6 with 2 and .4 with 1 at all steps of the solution below.

Answer: C: 0.24

Hint/Solution: A devilish problem—one of the most difficult actuarial exam problems I have seen! What makes this problem so tricky is the restriction to customers who made one claim in 1997 in the expectation to be computed. Without this restriction the expected number of claims would simply be an average of the expected numbers for high and low risk customers, weighted by the probabilities of high and low risk customers, i.e., $0.6 \cdot 0.1 + 0.1 \cdot 0.9 = 0.15$.

To take the above restriction into account, one has to replace the weights 0.1 and 0.9 in this calculation by corresponding *conditional* probabilities, the condition being that the customer had one accident in 1997. More formally, let H and L denote high and low risk customers, respectively, and let A denote the event "one claim made in 1997". Then the above weights are $P(H)(= 0.1)$ and $P(L)(= 0.9)$, and we need to replace these by $P(H|A)$ and $P(L|A)$, respectively.

The computation of these conditional probabilities is a nontrivial exercise in itself. By Bayes' rule we have

$$P(H|A) = \frac{P(A|H)P(H)}{P(A|H)P(H) + P(A|L)P(L)}.$$

By definition, $P(A|H)$ is the probability that (exactly) one claim is made, *given that the customer is high risk.* Since for high risk drivers the number of claims is Poisson distributed with mean 0.6, this probability is equal to $e^{-0.6}0.6$. Analogously, we get $P(A|L) = e^{-0.1}0.1$. Substituting these values along with $P(H) = 0.1$ and

15

Math 370 *A.J. Hildebrand* *Spring 2008*

$P(L) = 0.9$ into the above formula, we get $P(H|A) = 0.288$, and hence $P(L|A) = 1 - P(H|A) = 0.712$, by the complement formula for conditional probabilities. Using these weights in place of 0.1 and 0.9 in the above calculation, we get the correct answer: $0.6 \cdot 0.288 + 0.1 \cdot 0.712 = 0.244$.

Answers to Homework 3

1. 2/3
9. 2/3
12. (a) .27 (b) .675
24. .26
27. 2/5
32. .2/.72, .24/.72, .28/.72
40. 1/8
45. 4/5
48. (a) .07 (b) 3/7 (c) .09286
49. (a) 1/2 (b) 5/9
57. 4/7 and 1/7
64. would give too much away to give answer