Solutions 6

1. Raindrops are falling at an average rate of 20 drops per square inch per minute. What would be a reasonable distribution to use for the number of raindrops hitting a particular region measuring 5 $inches^2$ in t minutes? Why? Using your chosen distribution, compute the probability that the region has no rain drops in a given 3 second time interval.

   Solution: A reasonable choice of distribution is Poisson(λt), where λ= 20 · 5 = 100 (the average number of raindrops per minute hitting the region). Assuming this distribution,
   P(no raindrops in 1/20 of a minute) = $e^{-\frac{100}{20}}(\frac{100}{20})^0/0! = e^{-5}$.

2. Harvard Law School courses often have assigned seating to facilitate the "Socratic method." Suppose that there are 100 first year Harvard Law students, and each takes two courses: Torts and Contracts. Both are held in the same lecture hall (which has 100 seats), and the seating is uniformly random and independent for the two courses.

   (a) Find the probability that no one has the same seat for both courses (exactly; you should leave your answer as a sum).

   Solution: Let N be the number of students in the same seat for both classes. The problem has the same structure as the de Montmort matching problem from lecture. Let E_j be the event that the jth student sits in the same seat in both classes. Then

   $$P(N = 0) = 1 - P(\cup_{j=1}^{100} E_j)$$

   By symmetry, inclusion-exclusion gives

   $$P\left(\cup_{j=1}^{100} E_j\right) = \sum_{j=1}^{100}(-1)^{j-1}C_{100}^{j}P(\cap_{k=1}^{j} E_k)$$

   The j-term intersection event represents j particular students sitting pat throughout the two lectures, which occurs with probability (100 - j)!/100!. So

   $$P\left(\cup_{j=1}^{100} E_j\right) = \sum_{j=1}^{100}(-1)^{j-1}C_{100}^{j}\frac{(100-j)!}{100!} = \sum_{j=1}^{100}\frac{(-1)^{j-1}}{j!}$$

   $$P(n = 0) = 1 - \sum_{j=1}^{100}\frac{(-1)^{j-1}}{j!} = \sum_{j=1}^{100}\frac{(-1)^{j}}{j!}$$

   (b) Find a simple but accurate approximation to the probability that no one has the same seat for both courses.

   Solution: Define I_i to be the indicator for student i having the same seat in both courses, so that N = $\sum_{i=1}^{100} I_i$. Then P(I_i = 1) = 1/100, and the I_i are weakly dependent because
   P((I_i = 1) ∩ (I_j = 1)) = $C_{100}^{1}C_{99}^{1} \approx {C_{100}^{1}}^{2}$ = P(I_i = 1) P (I_j = 1)
   So N is close to Pois(λ) in distribution, where λ= E(N) = 100EI_i = 1. Thus,
   P(N = 0) ≈ $e^{-1}1^0$ /0! = e^{-1} .
   This agrees with the result of (a), which we recognize as the Taylor series for e^x, evaluated at x = - 1.

   (c) Find a simple but accurate approximation to the probability that at least two students have the same seat for both courses.

   Solution: Using a Poisson approximation, we have

$P(N \geq 2) = 1 - P(N = 0) - P(N = 1) \approx 1 - e^{-1} - e^{-1} = .1 - 2e^{-1}$

3. Explain why if X ~ Bin(n, p), then n - X ~ Bin(n, 1 - p).

   Solution: This follows immediately from the story of the Binomial, by redefining “success” as “failure” and vice versa.

4. There are 100 passengers lined up to board an airplane with 100 seats (with each seat assigned to one of the passengers). The first passenger in line crazily decides to sit in a randomly chosen seat (with all seats equally likely). Each subsequent passenger takes his or her assigned seat if available, and otherwise sits in a random available seat. What is the probability that the last passenger in line gets to sit in his or her assigned seat? (This is another common interview problem, and a beautiful example of the power of symmetry.)

   Hint: call the seat assigned to the jth passenger in line “Seat j” (regardless of whether the airline calls it seat 23A or whatever). What are the possibilities for which seats are available to the last passenger in line, and what is the probability of each of these possibilities?

   Solution: The seat for the last passenger is either Seat 1 or Seat 100; for example, Seat 42 can’t be available to the last passenger since the 42nd passenger in line would have sat there if possible. Seat 1 and Seat 100 are equally likely to be available to the last passenger, since the previous 99 passengers view these two seats symmetrically. So the probability that the last passenger gets Seat 100 is 1/2.

5. A stick is broken into two pieces, at a uniformly random chosen break point. Find the CDF and average of the length of the longer piece.

   Solution: We can assume the units are chosen so that the stick has length 1. Let L be the length of the longer piece, and let the break point be U ~ Unif(0, 1). For any l ∈ [1/2, 1], observe that L< l is equivalent to U < l, 1 – U < l, which can be written as 1 – l < U <l. We can thus obtain L’s CDF as
   $F_L(l) = P(L < l) = P(1 - l < U < l) = 2l - 1,$
   so L ~ Unif(1/2, 1) and E(L)=3/4.

6. Let U be a Uniform r.v. on the interval (-1, 1) (be careful about minus signs).

   (a) Compute E(U), Var(U), and E(U^4).

   Solution: We have E(U) = 0 since the distribution is symmetric about 0. By LOTUS,
   $$E(U^2) = \frac{1}{2}\int_{-1}^{1} u^2 d\,u = \frac{1}{3}$$
   So Var(U) = E(U^2) - $(EU)^2$ = E(U^2) = 1/3 . Again by LOTUS,
   $$E(U^4) = \frac{1}{2}\int_{-1}^{1} u^4 d\,u = \frac{1}{5}$$

   (b) Find the CDF and PDF of U^2. Is the distribution of U^2 Uniform on (0, 1)?

   Solution: Let G(t) be the CDF of U^2. Clearly G(t) = 0 for t ≤ 0 and G(t) = 1 for t 1, because $0 \leq U^2 \leq 1$. For 0 < t < 1,
   G(t) = $P(U^2 \leq t) = P(-\sqrt{t} \leq U \leq \sqrt{t}) = \sqrt{t},$

since the probability of U being in an interval in (-1, 1) is proportional to its length. The PDF is G' (t) = 1/2 $t^{-1/2}$ for 0 < t < 1 (and 0 otherwise). The distribution of U^2 is not Uniform on (0, 1) as the PDF is not a constant on this interval (it is an example of a Beta distribution, which is another important distribution in statistics and will be discussed later).

7. For X ~ Pois(λ), find E(X!) (the average factorial of X), if it is finite.

   Solution: By LOTUS,
   E(X!) = $e^{-\lambda} \sum_{k=0}^{\infty} k! \frac{\lambda^k}{k!} = \frac{e^{-\lambda}}{1-\lambda}$,
   for 0 < λ < 1 since this is a geometric series (and E(X!) is infinite if λ≥1).