Math 318 – homework 3 solutions

Problem 1. Two hockey teams, A and B play a series of games, until one of the teams wins 4 games. Suppose team A has probability p of winning each game, and games are independent. Let X be the total number of games that are played.

(a) Find the probability mass function of X.
(b) What is the probability that team A wins the series conditioned on $X = 4$?
(c) What is the probability that team A wins the series conditioned on $X = 7$? (Simplify your expressions as much as possible.)

Solution.

(a) Denote $q = 1 - p$ to simplify some formulas.

$$p(4) = p^4 + q^4 \qquad p(6) = 10p^4q^2 + 10p^2q^4$$
$$p(5) = 4p^4q + 4pq^4 \qquad p(7) = 20p^4q^3 + 20p^3q^4 = 20p^3q^3.$$

In each case the first term is the probability that A wins and the second that B wins in that many games. For example, to find $p(6)$, in order for team A to win in 6 games, they must lead 3-2 after 5 games ($\binom{5}{2} = 10$ ways for this to happen) and then win the last game. Each such sequence has probability p^4q^2.

(b) This is $\frac{p^4}{p^4+q^4}$.

(c) This is $\frac{20p^4q^3}{20p^3q^3} = p$. We can also see this directly, since if the series lasts 7 games, then after 6 games the score is 3-3, and the last game determines the winner.

Problem 2. A fair (6-sided) die is rolled four times.

(a) Let Y denote the number of distinct results. Find the probability mass function and expectation of Y.
(b) Let Z denote the minimal result out of the 4 throws. Find the probability mass function and expectation of Z.

Solution.

(a) We have

$$p(1) = \frac{6}{6^4} \qquad p(2) = \frac{14 \cdot \binom{6}{2}}{6^4} \qquad p(3) = \frac{\binom{6}{3} \cdot 3 \cdot 12}{6^4} \qquad p(4) = \frac{6 \cdot 5 \cdot 4 \cdot 3}{6^4}.$$

The cases 1 and 4 are easier. For $k = 2$, there are $\binom{6}{2}$ ways to choose which two values appear, and there are 14 sequences where the two specific values appear and no others (2^4 where no others appear, minus the two where only one of them appears). (The sequences are aaab,aaba,aabb,....) Similarly, if 3 values appear, there are $\binom{6}{3}$ ways to pick which three values appear, 3 ways to specify which of them appears twice, and 12 ways to arrange the throws with given outcomes $\{a, a, b, c\}$.

(b) We have $p(k) = \frac{(7-k)^4-(6-k)^4}{6^4}$ for $k = 1, 2, \ldots, 6$. This is since there are $(7 - k)^4$ sequences where the minimum is at least k, and we subtract $(6 - k)^4$ where the minimum is at least $k + 1$.

Problem 3. This problem investigates the similarity between the geometric and exponential random variables observed last week. Let Y be a geometric random variable with parameter p, so that Y represents the trial number of the first success in a sequence of independent Bernoulli trials. Suppose the trials occur at times $\delta, 2\delta, \ldots$, and that δ and p are both very small. Let $\lambda = p/\delta$. At time t, about t/δ trials have taken place.

(a) Compute $P(Y > m)$, which represents the probability that no success has been observed by time $t = m\delta$.

(b) Show that the probability that no success has been observed by time t converges to $e^{-\lambda t}$ as $p, \delta \to 0$ with $\lambda = p/\delta$ fixed.
(c) Conclude that the time of the first success is approximately an exponential random variable with parameter λ.

Solution.
(a) For integer m, $P(Y > m) = (1-p)^m$, since this is the probability of the first m experiments failing..
(b) by time t, the number of attempts is $[t/\delta]$, (here $[x]$ is the largest integer less than x). Therefore $P(Y > t) = (1-p)^{[t/\delta]} = e^{[t/\delta]\log(1-p)}$. As $p \to 0$ we have $\log(1-p) = -p + O(p^2)$, and therefore

$$\log P(Y > t) = -p[t/\delta] + O(p^2[t/\delta]).$$

The first term tends to $-\lambda t$ and the second to 0. Therefore $P(Y > t) \to e^{-\lambda t}$ as $p, \delta \to 0$.
(c) Since the CDF of Y is $1 - P(Y > t)$ and tends to $1 - e^{-\lambda t}$, which is the CDF of the exponential, Y is approximately exponential. We will discuss convergence of distributions in detail later this term.

Problem 4. A binary message either 0 or 1 is transmitted by wire. However, data sent over the wire is subject to channel noise disturbance. If x is the value sent (either 0 or 1), then the value received at the other end is $R = x + N$, where N represents the noise. Assume that N is a normal random variable with mean $\mu = 0$ and variance $\sigma^2 = 0.04$. Assume that a message sent is equally likely to be 0 or 1. When the message is received the receiver decodes it according to the following rule: If $R \leq \frac{1}{2}$ she concludes the message is 0, and otherwise concludes it is 1. What is the probability that the message is received correctly?

Solution. If the message is 0 we need $N < 1/2$. If the message is 1 we need $N > -0.5$. The probability is the same in both cases, and is $F(0.5)$, where F is the CDF Φ of the $N(0, 0.2^2)$ variable. Since $N(0,\sigma^2) = \sigma N(0,1)$, we can write $F(x) = \Phi(x/0.2)$, where Φ is the standard normal CDF. The answer is therefore $\Phi(0.5/0.2) = \Phi(2.5) = 0.993\ldots$. Note: $\Phi(x) = \frac{1+erf(x/\sqrt{2})}{2}$. This can be calculated in python as follows.

```
>>> import scipy.stats
>>> scipy.stats.norm.cdf(2.5)
0.99379033467422384
>>> from math import erf
>>> (1 + erf(2.5/sqrt(2)))/2
0.9937903346742238
```

Problem 5. The number of murders in Gotham on any given week is assumed to be Poisson with unknown mean λ, with different weeks independent.
(a) We observe there were 2 murders one week. What is the value of λ for which this is most likely?
(b) We observe for a second week, and there is 1 murder. What is the value of λ for which this pair of observations is most likely?
(c) Generalize the above to observations $a_1, \ldots, a_k$ over k weeks.

Solution.
(a) The likelihood of observing 2 murders in a week is $e^{-\lambda}\frac{\lambda^2}{2}$. This tends to 0 at $0, \infty$, and is maximized when $\lambda = 2$.
(b) The likelihood of the two observations is

$$L(\lambda) = e^{-\lambda}\frac{\lambda^2}{2} \cdot e^{-\lambda}\frac{\lambda}{1} = e^{-2\lambda}\frac{\lambda^3}{2}.$$

This is maximized at $\lambda = \frac{3}{2}$.

(c) If the observations are $a_1, \dots, a_k$, then the likelyhood is

$$L(\lambda) = \prod_i e^{-\lambda} \frac{\lambda^{a_i}}{a_i!} = e^{-k\lambda} \frac{\lambda^S}{\prod_i a_i!},$$

where $S = \sum a_i$. This is maximized at $\lambda = \frac{1}{k} \sum a_i$, which is the average observation. Note that this is consistent with the idea of λ being the average rate of murders.

Problem 6. An airline books passengers for a flight on an airplane with 420 seats. From experience, the airline knows that each passenger has probability $p = \frac{1}{50}$ of missing the flight. Assume these events are independent. As such, the airline takes a risk and sells 430 tickets for the flight.

(a) Using python, compute the probability that more than 420 passenger show up.

(b) Use the Poisson approximation to compute an approximation to this probability.

(c) Simulate the number of no-shows for an overbooked flight 50000 times. (You can use a function that returns a binomial random variable.) Plot a histogram of the fraction of times there were k no-shows, and the Poisson p.m.f. on the same graph.

(d) Simulated the number of no-shows 50000 times, and define X_n = number of overfull flights in the first n simulated bookings. Then X_n/n is the running proportion of overbooked flights. Plot X_n/n. What happens to it as n gets large?
Note: When calculating X_n, do not simulate n new bookings for every n. Simulate 50000 flights, and then for every n, calculate X_n based on the first n of these.