IEOR 3106: Introduction to Operations Research: Stochastic Models

Fall 2013, Professor Whitt, Tuesday, October 22

Three Poisson Problems

1. compound Poisson process, §5.4.2
2. Poisson random measure or Poisson process in the plane, §5.3.1 and Exercise 5.94
3. multiple proofreaders, Exercise 5.62

1. Money Withdrawn from an ATM Machine

Customers arrive at an automated teller machine (ATM) at the times of a Poisson process with a rate of $\lambda = 10$ per hour. Suppose that the amount of money withdrawn on each transaction has a mean of \$30 and a standard deviation of \$20.

(a) Find the mean and variance of the total amount of dollars withdrawn in 8 hours.

(b) What is the approximate probability that the total amount of money withdrawn in the first 8 hours exceeds \$3,400?

(c) How do the answers change if the Poisson arrival process is a *nonhomogeneous* Poisson process witharrival rate function $\lambda(t) = 2.5t$, $t \geq 0$?

Answers to Question 1.

(a) Assuming that successive withdrawals are IID, this is a **compound Poisson process**; see Section 5.4.2. Let $W(t)$ be the total amount withdrawn in the time interval $[0, t]$. Let $N(t)$ be the number of customers to come to the ATM in the interval $[0, t]$. Let Y_n be the amount of the n^{th} withdrawal. Then $W(t)$ can be represented as the following random sum of random variables

$$W(t) = \sum_{i=1}^{N(t)} Y_i \ .$$

Hence, from §5.4 (p. 346 in the last edition),

$$E[W(t)] = \lambda t E[Y_1] \quad \text{and} \quad var(X(t)) = \lambda t E[Y_1^2] \ , \tag{1}$$

so that

$$E[W(8)] = 10 \times 8 \times 30 = 2400 \quad \text{and} \quad var(W(8)) = 10 \times 8 \times ((30^2 + (20)^2) = 104,000 \ .$$

The standard deviation is $\sqrt{104,000} \approx 322.49$.

But why are those the correct formulas? To see why, look at Examples 3.11 and 3.19 in Chapter 3. The idea is to use conditioning. First, we use the formula $E[Y] = E[E[Y|Z]]$. Assuming that Z is integer-values, this is shorthand for

$$E[Y] = E[E[Y|Z]] = \sum_{n=1}^{\infty} E[Y|Z = n]P(Z = n).$$

If instead Z has a density $f_Z(z)$, we have

$$E[Y] = E[E[Y|Z]] = \int_{-\infty}^{\infty} E[Y|Z = t] f_Z(t)\, dt.$$

In our case, we look at

$$\begin{aligned} E[W(t)] &= E[\sum_{i=1}^{N(t)} Y_i] = \sum_{n=1}^{\infty} E[\sum_{i=1}^{N(t)} Y_i | N(t) = n] P(N(t) = n) \\ &= \sum_{n=1}^{\infty} n E[Y_1] p(N(t) = n) = E[N(t)] E[Y_1]. \end{aligned}$$

We next discuss the harder variance formula. Let

$$W = \sum_{i=1}^{N} Y_i \ ,$$

where N is a nonnegative-integer-valued random variable and Y_i are IID random variables. Then, by the conditional variance formula in Proposition 3.1,

$$Var(W) = E[Var(W|N)] + Var(E[W|N]) = E[N]Var(Y_1) + E[Y_1]^2 Var(N) \ .$$

In the Poisson-process case, with rate λ,

$$E[N(t)] = Var(N(t)) = \lambda t \ .$$

That yields formula (1) above.

(b) Use a normal approximation. It can be justified by applying the central limit theorem, because the stochastic process $\{W(t) : t \geq 0\}$ has stationary and independent increments and the summands all have finite mean and variance. We can think of

$$W(t) = \sum_{i=1}^{n} [W(k/n) - W((k-1)/n)],$$

which is the sum of i.i.d. random variables. Hence,

$$\begin{aligned} P(W(t) > 3400) &= P\left(\frac{W(t) - E[W(t)]}{std(W(t))} > \frac{3400 - E[W(t)]}{std(W(t))}\right) \\ &\approx P\left(N(0,1) > \frac{3400 - E[W(t)]}{std(W(t))}\right) \\ &\approx P\left(N(0,1) > \frac{3400 - 2400}{322}\right) = P\left(N(0,1) > \frac{1000}{322}\right) \\ &\approx P\left(N(0,1) > 3\right) \approx 0.0013. \end{aligned} \tag{2}$$

(c) The nonhomogeneous Poisson process is discussed in §5.4.1. The random variable $N(t)$ is still a Poisson random variable, but we need to calculate the new mean. In formulas (1), replace $\lambda t = 10 \times 8 = 80$ by

$$m(8) = \int_0^8 \lambda(s)\, ds = \int_0^8 2.5 s\, ds = 80.$$

Hence, the mean $E[N(t)]$ turns out to be unchanged. So the answers are all the same.

---

2. Random Art

A probabilistic painter decides to paint a large wall by a random process. He puts points on the wall according to a two-dimensional Poisson process (or Poisson random measure) with constant intensity (rate) 2 points per square foot.

(a) What is the probability that there are no points in a particular rectangle on the wall, which is 3 feet wide and 2 feet high?

(b) What is the probability that two disjoint 3×2 rectangles of the kind considered in part (a) each have exactly 10 points?

(c) Consider an arbitrary fixed position on the wall. What is the probability that the distance from that location to the nearest random point on the wall is greater than 6 inches (1/2 foot)?

This problem relates to Problem 5.94 in the textbook.

Answers to Question 2.

(a) The meaning of the model is that: (i) the number of points in any set has a Poisson distribution with a mean equal to the rate λ multiplied by the area of the set, and (ii) the numbers of points in disjoint sets are independent random variables. Let $N(C)$ be the number of points in this set (rectangle) C. Since the rate has been given to be 2 per square foot and the area of C is $2 \times 3 = 6$ square feet, the mean $E[N(C)] = 2 \times 6 = 12$.

$$P(N(C) = 0) = e^{-2 \times 6} = e^{-12},$$

not very likely.

(b) Because of the independence, we square the probability for each rectangle. (We can use the mean computed from the last part.) The answer is

$$\left(\frac{e^{-12}(12)^{10}}{10!} \right)^2$$

(c) This is equivalent to there being no points in the set C, where now C is the circle of radius 1/2 foot. The circle of radius 1/2 has area $\pi r^2 = \pi/4 \approx 0.785$. Hence,

$$P(N(C) = 0) = e^{-\lambda \pi r^2} = e^{-2 \times 0.785} = e^{-1.570} \approx 0.208$$

---

3. Typographical Errors, Exercise 5.62 on p. 363.

Suppose that the number of typographical errors in a new text is Poisson distributed with mean λ.

We can arrive at this model by considering a more detailed model: More specifically, suppose that the text is 100 pages with 100 lines per page and 100 character spaces per line, yielding $10^6 = 1,000,000$ character spaces in all. As a rough approximation, we may regard the occurrence of errors in specific characters as being independent and identically distributed Bernoulli random variables. The total number of errors in any subset of text would then have a binomial distribution (exactly). However, we can use the Poisson approximation for the binomial distribution, which becomes more and more appropriate as n increases and p decreases. The Poisson distribution is given mean $\lambda = np$.. With that approximation, we might regard the number of errors in various subsets of the text as a *Poisson random measure*. Suppose that A is a portion of text containing k character spaces. The mean number of errors the subset A of the text would be

$$E[N(A)] = \frac{\lambda \times k}{10^6} .$$

Then the expected total number of errors in the entire text is simply λ.

Now suppose that two proofreaders independently read the text. Suppose that each error is independently found by proofreader i with probability p_i, $i = 1, 2$. Let X_1 be the number of errors found by proofreader 1, but not proofreader 2; Let X_2 be the number of errors found by proofreader 2, but not proofreader 1. Let X_3 be the number of errors that are found by both proofreaders; and let X_4 be the number of errors found by neither proofreader.

The first goal is to find an estimator for the distribution of X_4.

The second goal is to estimate the benefit of having a new third proofreader read the text, where this proofreader finds each error independently with probability q.

(a) Describe the joint distribution of X_1, X_2, X_3 and X_4.

(b) Verify that

$$\frac{E[X_1]}{E[X_3]} = \frac{1 - p_2}{p_2} \quad \text{and} \quad \frac{E[X_2]}{E[X_3]} = \frac{1 - p_1}{p_1} .$$

(c) By using X_i as an estimator of $E[X_i]$, present estimators of p_1, p_2 and λ.

(d) Give an estimate of $E[X_4]$ the expected number of errors not found by either proofreader.

(e) Suppose $X_1 = 60$, $X_2 = 30$ and $X_3 = 40$. What is the estimated distribution of X_4?

(f) Now we contemplate hiring the third proofreader who independently finds errors with probability $q = .9$. How much do we reduce the expected number of uncovered errors by using this third proof reader.

Answers to Question 3.

(a) Here we use the basic independent thinning property for Poisson random variables and Poisson processes; see §5.3.4 and Proposition 5.2. The random variables X_1, X_2,

X_3 and X_4 are mutually independent Poisson random variables with means

$$\begin{aligned} E[X_1] &= \lambda p_1(1-p_2) \\ E[X_2] &= \lambda p_2(1-p_1) \\ E[X_3] &= \lambda p_1 p_2 \\ E[X_4] &= \lambda(1-p_1)(1-p_2) \ . \end{aligned}$$

Note, however, that we do not observe X_4, but we do observe X_1, X_2 and X_3. From above, we know that the exact distribution of X_4 is Poisson, but with a mean $E[X_4] = \lambda(1-p_1)(1-p_2)$. However, we do not know the three parameters λ, p_1 and p_2 appearing in that expression. The idea is to estimate $E[X_4]$ by estimating these three parameters. The idea is to use the three quantities observed (X_1, X_2 and X_3) to estimate the three unknown parameters (λ, p_1 and p_2). We start by estimating the three means $E[X_1]$, $E[X_2]$ and $E[X_3]$ by the corresponding observed values X_1, X_2 and X_3. After that step, it remains only to solve the three equations above in the three unknowns

---

(b) Follows easily from part (a).

(c) First, by (b),

$$p_1 = \frac{1}{1 + \frac{E[X_2]}{E[X_3]}} = \frac{E[X_3]}{(E[X_2] + E[X_3])} \ .$$

Hence we can estimate p_1 by $X_3/(X_2+X_3)$. Thus p_1 is estimated by the fraction of error found by proofreader 2 that are also found by proofreader 1. Similarly (just change the labels!), we can estimate p_2 by $X_3/(X_1+X_3)$.

The total number of errors found has mean

$$E[X_1 + X_2 + X_3] = E[X_1] + E[X_2] + E[X_3] = \lambda(1 - (1-p_1)(1-p_2)) \ ,$$

so that

$$E[X_1 + X_2 + X_3] = \lambda(1 - (1-p_1)(1-p_2)) = \lambda(1 - \frac{E[X_2]E[X_1]}{(E[X_2] + E[X_3])(E[X_1] + E[X_3])} \ .$$

Hence we can estimate λ by

$$\hat{\lambda} = \frac{(X_1 + X_2 + X_3)}{(1 - \frac{X_2 X_1}{(X_2+X_3)(X_1+X_3)})} \ .$$

(d) Note that

$$E[X_4] = \lambda - (E[X_1] + E[X_2] + E[X_3]) \ ,$$

so we can estimate $E[X_4]$ by

$$\hat{\lambda} - (X_1 + X_2 + X_3) = (X_1 + X_2 + X_3)(\frac{Z}{(1-Z)}) \ ,$$

where

$$Z \equiv \frac{X_2 X_1}{(X_2 + X_3)(X_1 + X_3)} \ .$$

The above amounts to explicitly solving the three equations in the three unknowns.

(e) As stated above, we know that X_4 has a Poisson distribution. We estimate its mean $E[X_4]$ by

$$130(\frac{(18/70)}{1 - (18/70)} = 130\frac{18}{52} = 45 \ .$$

We have just plugged in the numbers to the formula already developed.

(f) The number of undiscovered errors before using this new proofreader is X_4. The number of undiscovered errors after using this new proofreader is Poisson distributed with mean $E[X_4](1 - q)$. We anticipate that the a proportion $(1 - q)$ of the remaining errors will remain. We will expect to delete $E[X_4]q$ errors. That is, we will discover a proportion q of the remaining errors if we use the third proof reader. We expect to reduce the number of undiscovered errors from 45 to 4.5.
