IEOR 3106: Introduction to Operations Research: Stochastic Models

Professor Whitt, Fall 2013

Solutions to Problem for Discussion, Thursday, October 10

Gone Fishing

To seek a change of pace from his intense Columbia experience, Seth Hochhauser has decided to go fishing. It is conceivable that this, like everything else, could be done in a brief excursion from campus, but Seth wants a real change. So Seth has decided to go to Venice, Florida, which is centrally located on the Gulf of Mexico along the coast of southwest Florida. As usual, he hopes to catch some grouper and snapper, but she is also hoping to catch some other fish, such as kingfish, cobia, black-fin tuna, Greater amberjack, Spanish mackerel, dolphin fish (mahi-mahi, not dolphins), shark, barracuda, tarpon, permit, little tunny, sheepshead, flounder, snook, redfish, and sea trout.

Suppose that Seth catches fish according to a Poisson process at a rate of 3 per hour.

(a) What is the expected number of fish he catches in a two-hour period?

Let $N(t)$ be the number of fish Seth catches in the time interval $(0, t)$, where time is measured in hours. The assumption is that the stochastic process $\{N(t) : t \geq 0\}$ is a Poisson process. One implication is that $N(t + s) - N(s)$ has a Poisson distribution with mean $3t$ for any $t > 0$ and $s > 0$. It suffices to focus on $N(t) = N(t) - N(0)$, where $N(0) = 0$. The distribution is

$$P(N(t) = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!} \ ,$$

where λ is the rate and t is the length of the time interval. In other words, the random variable $N(t)$ has a Poisson distribution with mean λt.

Here $\lambda = 3$ and $t = 2$. So the distribution is

$$P(N(2) = k) = \frac{e^{-3\times 2}(3 \times 2)^k}{k!} = \frac{e^{-6}(6)^k}{k!}$$

Now, answering the specific question, recalling that the mean of a Poisson distribution with parameter λ is just λ, so that $E[N(t)] = \lambda t$ and

$$E[N(2)] = \lambda t = 3 \times 2 = 6$$

(b) What is the variance of the number of fish he catches in a two-hour period?

For a Poisson distribution, the variance equals the mean. Thus,

$$Var(N(2)) = E[N(2)] = 6$$

using part (a)

(c) What is the probability that he catches exactly 4 fish in a given 2-hour period?

From above,

$$P(N(2) = 4) = \frac{e^{-3\times 2}(3 \times 2)^4}{4!} = \frac{e^{-6}(6)^4}{4!} = 0.134$$

(d) What is the probability that he catches at least two fish in a two-hour period?

$$\begin{aligned} P(N(2) \geq 2) = 1 - P(N(2) = 0) - P(N(2) = 1) &= 1 - \frac{e^{-6}(6)^0}{0!} - \frac{e^{-6}(6)^1}{1!} \\ &= 1 - e^{-6} - 6e^{-6} = 1 - 7e^{-6} = 0.8785 \end{aligned}$$

(e) What is the *conditional* probability that he catches exactly 4 fish in a given 2-hour period, given that he catches 23 fish in the previous two hours?

A fundamental property of the Poisson process is that it has stationary increments and independent increments. Stationary increments means that the distribution of $N(t+s) - N(s)$ is independent of s. Independent increments means that $N(s+t) - N(s)$ is independent of $N(a+b) - N(a)$ for nonnegative real numbers s, t, a and b, provided that the two intervals $(s, s+t)$ and $(a, a+b)$ are disjoint. Hence, the conditioning makes no difference (according to the model!) So

$$P(N(t+2) - N(t) = 4|N(t) - N(t-2) = 23) = P(N(t+2) - N(t) = 4) = \frac{e^{-6}(6)^4}{4!} = 0.134 ,$$

just as in part (c) above.

(f) What is the *conditional* expected number of fish that he catches in a given 2-hour period, given that he catches 23 fish in the previous two hours?

By the independence mentioned above in part (e),

$$E[N(t+2) - N(t)|N(t) - N(t-2) = 23] = E[N(t+2) - N(t)] = 6 ,$$

just as in part (a).

(g) Given your answers to parts (e) and (f), what feature of the probability model would you question? How could you test whether or not that model feature is reasonable for this kind of fishing?

For fishing, we might not believe the independent-increments property. The Poisson distribution might be a reasonable model for one given time interval, but we might not assume independent increments.

Testing whether independent increments holds is a standard statistical problem, but not always easy in applications. The general idea is to measure the numbers of fish caught in disjoint intervals, and then see if the results are consistent with independence. The general idea is the same as the problem 2 (d) on the posted 2011 first midterm exam. But the specific techniques would be different in this problem. The problem is similar to problems arising in finance: In finance, stock prices are often modelled as geometric Brownian motion. That means that the logarithm of the stock price is modeled as ordinary Brownian motion. That implies that the logarithm of stock price should have independent increments. (Brownian motion is like the Poisson process in that both have independent increments, but Brownian motion has continuous sample paths, whereas the Poisson process has integer-values sample paths.) It is natural to want to test that. We do not try to provide details here; the question is beyond the scope of the course.

A rough test, with a lot of data, would be to see if the variance of $N(t+4) - N(t)$ is different from the sum of the variances of $N(t+4) - N(t+2)$ and $N(t+2) - N(t)$. In general, letting Var be the variance and Cov the covariance,

$$\begin{aligned}
&Var(N(t+4) - N(t)) = Var(N(t+4) - N(t+2) + N(t+2) - N(t)) \\
&= Var(N(t+4) - N(t+2)) + Var(N(t+2) - N(t)) \\
&\qquad +2Cov(N(t+2) - N(t), N(t+4) - N(t+2))
\end{aligned}$$

Hence,

$$\begin{aligned}
&Var(N(t+4) - N(t)) - [Var(N(t+4) - N(t+2)) + Var(N(t+2) - N(t))] \\
&\quad = 2Cov(N(t+2) - N(t), N(t+4) - N(t+2)).
\end{aligned}$$

So a difference would indicate a non-zero covariance, violating independence. Of course, even with independence, we cannot expect to see exactly 0 covariance in an estimate because of the randomness. Careful statistical analysis investigates how much deviation could not have happened by chance.

(h) What is the expected time until he catches his *fourth* fish?

For a Poisson process, the intervals between successive events are IID exponential random variables, with a mean equal to the reciprocal of the rate. So here the mean length of the interval between successive catching of fish is $1/\lambda = 1/3$ hour $= 20$ minutes. Let X_i be the length of the interval between catching his $(i-1)^{st}$ and i^{th} fish. Then we want

$$E[X_1 + X_2 + X_3 + X_4] = 4E[X_1] = \frac{4}{3} \ .$$

(i) What is the variance of the time until he catches his *fourth* fish?

Because of the IID property,

$$Var[X_1 + X_2 + X_3 + X_4] = 4Var[X_1] = 4E[X_1]^2 = \frac{4}{9} \ .$$

(j) How likely would it be that he would catch more than 90 fish in 27 hours?

Here we can use the normal approximation for the Poisson distribution. That can be based on the central limit theorem because $N(27) \stackrel{d}{=} N_1(1) + N_2(1) + \cdots + N_{27}(1)$, where $N_i(1)$ is the number of fish caught in the i^{th} hour, i.e., $N_i(1) \stackrel{d}{=} N(i) - N(i-1)$. At any rate, we can apply a normal approximation. The mean is $E[N(27)] = 3 \times 27 = 81$. Here, the variance is equal to the mean, so the standard deviation is equal to 9.

$$P(N(27) > 90) \approx P(\frac{N(27) - 81}{9} > \frac{90 - 81}{9}) \approx P(N(0,1) > 1) \approx 0.16$$

(k) How likely would it be that it would take him less than 27 hours to catch 90 fish?

Here we can use the normal approximation for sums of IID random variables. As it should appear, this problem is related to the previous one, being sort of an inverse problem. The time required to catch 90 fish is the sum of 90 IID exponential random variables, each with mean 1/3, and thus variance 1/9. The total mean is 30, while the total variance is 10, so the total standard deviation is 3.162. Let S_n be the sum of n IID exponentials. Hence,

$$P(S_{90} < 27) = P(\frac{S_{90} - 30}{3.162} < \frac{27 - 30}{3.162}) \approx P(N(0,1) < -\frac{3}{3.162}) \approx P(N(0,1) < -1) \approx 0.16$$

The important inverse relation is

$$S_n \le t \quad \text{if and only if} \quad N(t) \ge n \ ;$$

see the middle of page 294 in Ross. Thus, we could, with some work, deduce the answer here from the answer to the previous part.

(l) Suppose that he catches exactly 4 fish in a given 2-hour period. What then is the probability that he catches all four fish in the first 30 minutes?

We here apply Section 5.3.5 of Ross. Conditional on a Poisson process having n events in an interval (a, b), those n events are distributed throughout the interval (a, b) as n IID random variables, each distributed uniformly over the interval (a, b). The time to the first point, is thus the minimum of n IID uniform random variables.

For each of the four fish caught, the probability that it was caught in the first half hour is thus 1/4. The answer here then is $(1/4)^4 = 1/256 \approx 0.004$.

(m) Suppose that he catches exactly 4 fish in a given 2-hour period. What then is the probability that he catches his first fish after 30 minutes?

He catches his first fish after 30 minutes if and only if he catches all four fish after 30 minutes. So the probability is $(3/4)^4 = 81/256 \approx 0.316$

Suppose, in addition, that each fish Seth catches is a grouper with probability $1/4$, a snapper with probability $1/3$ and some other kind of fish with probability $5/12$, with the successive kinds being independent random trials.

(n) What is the probability that he catches exactly 8 fish in a given 2-hour period, with 3 of them being grouper and 5 being snapper?

Here we have the independent-splitting property. Let $N_G(t)$ be the number of grouper caught in the interval $(0, t)$; let $N_S(t)$ be the number of snapper caught in the interval $(0, t)$; let $N_O(t)$ be the number of other fish caught in the interval $(0, t)$. The key fact is that these three stochastic processes are independent Poisson processes. The rate for the grouper is the overall rate 3 times the probability $1/4$, which equals $3/4$. Thus the answer is the product of three Poisson probabilities:

$$\begin{aligned} P(N_G(2) = 3, N_S(2) = 5, N_0(2) = 0) &= P(N_G(2) = 3)P(N_S(2) = 5)P(N_0(2) = 0) \\ &= \frac{e^{-1.5}(1.5)^3}{3!}\frac{e^{-2}(2)^5}{5!}\frac{e^{-2.5}(2.5)^0}{0!} \\ &= \frac{e^{-6.0}(1.5)^3(2)^5}{3!5!} = \frac{(0.002479)(3.375)(32)}{720} = 0.000372 \end{aligned}$$

(Be sure to include that the other kin must be 0.)

(o) What is the conditional probability that he catches 8 grouper in one 2-hour period, given that he later catches 3 snapper in a subsequent two-hour period (according to the model)?

There is independence, because these are two independent Poisson processes. So

$$P(N_G(2) = 8 | N_S(t + 2) - N_S(t) = 3) = P(N_G(2) = 8) = \frac{e^{-1.5}(1.5)^8}{8!}$$

(p) What is the *conditional* probability that he catches 8 grouper in a given 2-hour period, given that he catches 14 snapper in the *same* two-hour period?

The answer is the same, because there still is independence:

$$P(N_G(2) = 8 | N_S(2) = 14) = P(N_G(2) = 8) = \frac{e^{-1.5}(1.5)^8}{8!}$$
