## IEOR 3106: Fall 2013, Professor Whitt

## Poisson Entertainment, Thursday, October 10

1. **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 he 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 he 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 that he catches in a two-hour period?

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

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

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

(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?

(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?

(g) Given your answers to parts (e) and (f), what feature of the probability model would you question?

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

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

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

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

(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?

(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?

Suppose, in addition, that each fish he 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?

(o) What is the probability that he catches 8 grouper in one 2-hour period, and then later catches 3 snapper in a subsequent two-hour period?

(p) What is the *conditional* probability that he catches 3 grouper in a given 2-hour period, given that he catches 14 snapper in the same two-hour period?

## 2. Technical background

### 2.1. A Point Process and a Counting Process

A **point process** on the positive half line, i.e., on the interval $[0, \infty)$, is a random distribution of points on the positive half line. We may specify the distribution in three ways: (i) by specifying the distribution of the locations of the points, (ii) by specifying the distribution of the intervals between successive points and (iii) by specifying the distribution of the associated counting process. Let $S_n$ be the location of the $n^{\text{th}}$ point, where $S_0 \equiv 0$ (without there being a $0^{\text{th}}$ point). Let $X_n \equiv S_n - S_{n-1}$ be the interval between the $(n-1)^{\text{st}}$ point and the $n^{\text{th}}$ point. Let the associated **counting process** be defined by

$$N(t) \equiv \max \{k \geq 0 : S_k \leq t\}, \quad t \geq 0.$$

In other words, a point process may be specified in three ways, via the stochastic processes: (i) $\{S_n : n \geq 0\}$, (ii) $\{X_n : n \geq 1\}$ and (iii) $\{N(t) : t \geq 0\}$. The first representation $\{S_n : n \geq 0\}$ is the typical form for a **point process**. The last representation $\{N(t) : t \geq 0\}$ is the typical form for a **counting process**.

A picture makes this clear; see Figure 1.

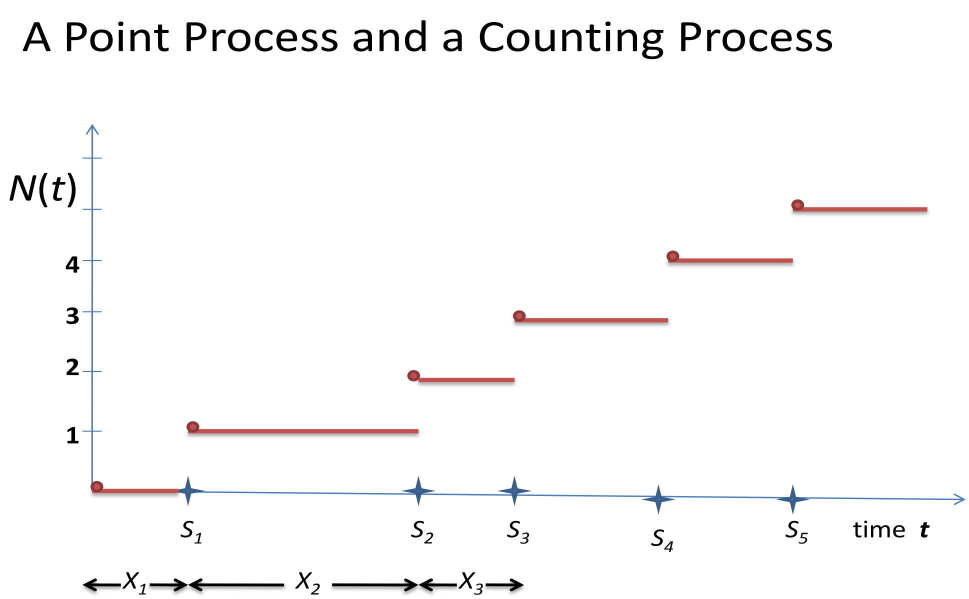


Figure 1: A Sample Path of a Counting Process.

For any point process or counting process, there is an important **inverse relation**, mentioned in Section 5.3.3 after Proposition 5.1 and discussed at greater length in (7.2) of Chapter

7. For any nonnegative $n$ and $t$,

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

This is for any possible realization; i.e., it is valid with probability 1. Again, a picture makes this clear: In Figure 1, the counting process view looks at the horizontal $x$ axis as the domain and the vertical $y$ axis as the range (mapping time $t$ into the number $N(t)$), while the point process view looks at the vertical $y$ axis as the domain and the horizontal $x$ axis as the range (mapping the nonnegative integers $n$ into the $n^{\text{th}}$ point $S_n$).

### 2.2. Alternative Definitions of a Poisson Process

(a) **Standard Definition**

The standard definition of a Poisson process is a counting process $\{N(t) : t \ge 0\}$ such that (i) $N(t)$ has a **Poisson distribution** with mean $\lambda t$ for each $t > 0$, where $\lambda > 0$ is some parameter, and (ii) the process has **independent increments**.

Property (i) means that

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

for any nonnegative integer $k$ (where $x^0 = 1$ and $0! = 1$). As a consequence,

$$E[N(t)] = \text{Variance}(N(t)) = \lambda t.$$

The stochastic process $N \equiv \{N(t) : t \ge 0\}$ has **independent increments** if the number of point in any number of **disjoint intervals** are independent random variables. The number of points in the interval $(a, b]$, closed on the right and open on the left, is $N(b) - N(a)$. (The probability of a point at any specific location will be 0, because the distance between successive points has a density.) An **increment** of the stochastic process $N \equiv \{N(t) : t \ge 0\}$ is $N(b) - N(a)$.

Suppose that $(t_1, t_2]$, $(t_3, t_4] \ldots (t_{2k-1}, t_{2k}]$ are $k$ disjoint intervals, i.e., with

$$0 \le t_1 < t_2 \le t_3 < t_4 \le \cdots \le t_{2k-1} < t_{2k}.$$

Then the $k$ random variables $N(t_2) - N(t_1)$, $N(t_4) - N(t_3), \ldots N(t_{2k}) - N(t_{2k-1})$ are mutually independent random variables. As a consequence, if $0 \le t_1 < t_2 \le t_3 < t_4$, then

$$\begin{aligned} P(N(t_2) - N(t_1) = j, N(t_4) - N(t_3) = k) &= P(N(t_2) - N(t_1) = j)P(N(t_4) - N(t_3) = k) \\ &= \frac{e^{-\lambda(t_2 - t_1)}(\lambda(t_2 - t_1))^j}{j!} \frac{e^{-\lambda(t_4 - t_3)}(\lambda(t_4 - t_3))^k}{k!} \end{aligned}$$

(b) **renewal process**

A point process (or counting process) is a renewal process if the intervals between points, i.e., the random variables $X_n$ defined above, are independent and identically distributed (i.i.d.). In renewal theory (Chapter 7), much attention is given to the associated renewal counting process $\{N(t) : t \ge 0\}$.

A Poisson process (Chapter 5) is a special case of a renewal process (Chapter 7) in which the times between renewals have an exponential distribution. This corresponds to Proposition 5.1 in Section 5.3.

Let us compute the distribution of $X_1$:

$$P(X_1 > t) = P(N(t) = 0) = \frac{e^{-\lambda t}(\lambda t)^0}{0!} = e^{-\lambda t}, \quad t \geq 0.$$

Hence, the cumulative distribution function (cdf) of $X_1$ is

$$F_{X_1}(t) \equiv P(X_1 \leq t) = 1 - e^{-\lambda t} \quad \text{(i.e., exponential with mean } 1/\lambda)$$

and the associated probability density function (pdf) is

$$f_{X_1}(t) = \lambda e^{-\lambda t}, \quad t \geq 0.$$