POISSON PROCESSES

1. THE LAW OF SMALL NUMBERS

1.1. **The Rutherford-Chadwick-Ellis Experiment.** About 90 years ago Ernest Rutherford and his collaborators at the Cavendish Laboratory in Cambridge conducted a series of pathbreaking experiments on radioactive decay. In one of these, a radioactive substance was observed in $N = 2608$ time intervals of 7.5 seconds each, and the number of decay particles reaching a counter during each period was recorded. The table below shows the number N_k of these time periods in which exactly k decays were observed for $k = 0, 1, 2, \dots, 9$. Also shown is Np_k where

$$p_k = (3.87)^k \exp\{-3.87\}/k!$$

The parameter value 3.87 was chosen because it is the mean number of decays/period for Rutherford's data.

k	N_k	Np_k	k	N_k	Np_k
0	57	54.4	6	273	253.8
1	203	210.5	7	139	140.3
2	383	407.4	8	45	67.9
3	525	525.5	9	27	29.2
4	532	508.4	≥ 10	16	17.1
5	408	393.5			

This is typical of what happens in many situations where counts of occurences of some sort are recorded: the Poisson distribution often provides an accurate – sometimes remarkably accurate – fit. Why?

1.2. **Poisson Approximation to the Binomial Distribution.** The ubiquity of the Poisson distribution in nature stems in large part from its connection to the Binomial and Hypergeometric distributions. The Binomial-(N, p) distribution is the distribution of the number of successes in N independent Bernoulli trials, each with success probability p. If p is small, then successes are *rare events*; but if N is correspondingly large, so that $\lambda = Np$ is of moderate size, then there are just enough trials so that a few successes are likely.

Theorem 1. *(Law of Small Numbers) If $N \to \infty$ and $p \to 0$ in such a way that $Np \to \lambda$, then the Binomial-(N, p) distribution converges to the Poisson-λ distribution, that is, for each $k = 0, 1, 2, \dots,$*

$$\binom{N}{k} p^k (1-p)^{N-k} \longrightarrow \frac{\lambda^{k e^{-\lambda}}}{k!} \tag{1}$$

What does this have to do with the Rutherford-Chadwick-Ellis experiment? The radioactive substances that Rutherford was studying are composed of large numbers of individuals atoms – typically on the order of $N = 10^{24}$. The chance that an individual atom will decay in a short time interval, and that the resulting decay particle will emerge in just the right direction so as to colide with the counter, is very small. Finally, the different atoms are at least approximately independent.[1]

Theorem 1 can be proved easily by showing that the probability generating functions of the binomial distributions converge to the generating function of the Poisson distribution. Alternatively, it is possible (and not very difficult) to show directly that the densities converge. For either proof, the following important analytic lemma is needed.

Lemma 2. *If λ_n is a sequence of real numbers such that $\lim_{n\to\infty} \lambda_n = \lambda$ exists and is finite, then*

$$\lim_{n\to\infty}\left(1-\frac{\lambda_n}{n}\right)^n = e^{-\lambda}. \tag{2}$$

Proof. Write the product on the left side as the exponential of a sum:

$$(1-\lambda_n/n)^n = \exp\{n\log(1-\lambda_n/n)\}.$$

This makes it apparent that what we must really show is that $n\log(1-\lambda_n/n)$ converges to $-\lambda$. Now λ_n/n converges to zero, so the log is being computed at an argument very near 1, where $\log 1 = 0$. But near $x = 1$, the log function is very well approximated by its tangent line, which has slope 1 (because the derivative of $\log x$ is $1/x$). Hence,

$$n\log(1-\lambda_n/n) \approx n(-\lambda_n/n) \approx -\lambda.$$

To turn this into a rigorous proof, use the second term in the Taylor series for log to estimate the error in the approximation. □

Exercise 1. Prove Theorem 1 either by showing that the generating functions converge or by showing that the densities converge.

The Law of Small Numbers is closely related to the next proposition, which shows that the exponential distribution is a limit of geometric distributions.

Proposition 3. *Let T_n be a sequence of geometric random variables with parameters p_n, that is,*

$$P\{T_n > k\} = (1-p_n)^k \quad \text{for } k = 0,1,2,\dots. \tag{3}$$

If $np_n \to \lambda > 0$ as $n \to \infty$ then T_n/n converges in distribution to the exponential distribution with parameter λ.

Proof. Set $\lambda_n = np_n$; then $\lambda_n \to \lambda$ as $n \to \infty$. Substitute λ_n/n for p_n in (3) and use Lemma 2 to get

$$\lim_{n\to\infty} P\{T_n > nt\} = (1-\lambda_n/n)^{[nt]} \longrightarrow \exp\{-\lambda t\}.$$

□

[1] This isn't completely true, though, and so observable deviations from the Poisson distribution will occur in larger experiments.

2. THINNING AND SUPERPOSITION PROPERTIES

Superposition Theorem. *If $Y_1, Y_2, \dots, Y_n$ are independent Poisson random variables with means $EY_i = \lambda_i$ then*

$$\sum_{i=1}^{n} Y_i \sim \text{Poisson}-\left(\sum_{i=1}^{n} \lambda_i\right) \tag{4}$$

There are various ways to prove this, none of them especially hard. For instance, you can use probability generating functions (Exercise.) Alternatively, you can do a direct calculation of the probability density when $n = 2$, and then induct on n. (Exercise.) But the clearest way to see that theorem theorem must be true is to use the Law of Small Numbers. Consider, for definiteness, the case $n = 2$. Consider independent Bernoulli trials X_i, with small successs parameter p. Let $N_1 = [\lambda_1/p]$ and $N_2 = [\lambda_2/p]$ (here $[x]$ means the integer part of x), and set $N = N_1 + N_2$. Clearly,

$$\sum_{i=1}^{N_1} X_i \sim \text{Binomial}-(N_1, p),$$
$$\sum_{i=N_1+1}^{N_2+N_1} X_i \sim \text{Binomial}-(N_2, p), \text{and}$$
$$\sum_{i=1}^{N} X_i \sim \text{Binomial}-(N, p).$$

The Law of Small Numbers implies that when p is small and N_1, N_2, and N are correspondingly large, the three sums above have distributions which are close to Poisson, with means λ_1, λ_2, and λ, respectively. Consequently, (4) has to be true when $n = 2$. It then follows by induction that it must be true for all n.

Thinning Theorem. *Suppose that $N \sim \text{Poisson}(\lambda)$, and that $X_1, X_2, \dots$ are independent, identically distributed Bernoulli-p random variables independent of N. Let $S_n = \sum_{i=1}^{n} X_i$. Then S_N has the Poisson distribution with mean λp.*

This is called the "Thinning Property" because, in effect, it says that if for each occurence counted in N you toss a p–coin, and then record only those occurences for which the coin toss is a Head, then you still end up with a Poisson random variable.

Proof. You can prove this directly, by evaluating $P\{S_N = k\}$ (exercise), or by using generating functions, or by using the Law of Small Numbers (exercise). The best proof is by the Law of Small Numbers, in my view, and I think it worth one's while to try to understand the result from this perspective. But for variety, here is a generating function proof: First, recall that the generating function of the binomial distribution is $Et^{S_n} = (q + pt)^n$, where $q = 1 - p$. The generating function of the Poisson distribution with parameter λ is

$$Et^N = \sum_{k=0}^{\infty} \frac{(t\lambda)^k}{k!} e^{-\lambda} = \exp\{\lambda t - \lambda\}.$$

Therefore,

$$\begin{aligned} E\,t^{S_N} &= \sum_{n=0}^{\infty} E\,t^{S_n} \lambda^n e^{-\lambda}/n! \\ &= \sum_{n=0}^{\infty} (q+pt)^n \lambda^n e^{-\lambda}/n! \\ &= \exp\{(\lambda q + \lambda p t) - \lambda\} \\ &= \exp\{\lambda p t - \lambda p\}, \end{aligned}$$

and so it follows that S_N has the Poisson distribution with parameter λp. □

A similar argument can be used to prove the following generalizaton:

Generalized Thinning Theorem . *Suppose that* $N \sim \text{Poisson}(\lambda)$*, and that* $X_1, X_2, \dots$ *are independent, identically distributed* multinomial *random variables with distribution* Multinomial—$(p_1, p_2, \dots, p_m)$*, that is,*

$$P\{X_i = k\} = p_k \quad \textit{for each } k = 1, 2, \dots, m$$

Then the random variables $N_1, N_2, \dots, N_m$ *defined by*

$$N_k = \sum_{i=1}^{N} \mathbf{1}\{X_i = k\}$$

are independent Poisson random variables with parameters $E N_k = \lambda p_k$.

Notation: The notation $\mathbf{1}_A$ (or $\mathbf{1}\{\dots\}$) denotes the *indicator function* of the event A (or the event $\{\dots\}$), that is, the random variable that takes the value 1 when A occurs and 0 otherwise. Thus, in the statement of the theorem, N_k is the number of multinomial trials X_i among the first N that result in outcome k.

3. Poisson Processes

Definition 1. A *point process* on the timeline $[0, \infty)$ is a mapping $J \mapsto N_J = N(J)$ that assigns to each subset[2] $J \subset [0, \infty)$ a nonnegative, integer-valued random variable N_J in such a way that if $J_1, J_2, ..$ are pairwise disjoint then

$$N(\cup_i J_i) = \sum_i N(J_i) \tag{5}$$

The *counting process* associated with the point process $N(J)$ is the family of random variables $\{N_t = N(t)\}_{t \geq 0}$ defined by

$$N(t) = N((0, t]). \tag{6}$$

Note: The sample paths of $N(t)$ are, by convention, always *right*-continuous, that is, $N(t) = \lim_{\varepsilon \to 0+} N(t + \varepsilon)$.

There is a slight risk of confusion in using the same letter N for both the point process and the counting process, but it isn't worth using a different letter for the two. The counting process $N(t)$ has sample paths that are step functions, with jumps of integer sizes. The discontinuities represent *occurences* in the point process.

[2] actually, to each *Borel measurable* subset

Definition 2. A *Poisson point process* of intensity $\lambda > 0$ is a point process $N(J)$ with the following properties:

(A) If $J_1, J_2, \dots$ are nonoverlapping intervals of $[0,\infty)$ then the random variables $N(J_1), N(J_2), \dots$ are mutually independent.
(B) For every interval J, the random variable $N(J)$ has the Poisson distribution with mean $\lambda|J|$, where $|J|$ is the length of J.

The counting process associated to a Poisson point process is called a *Poisson counting process*. Property (A) is called the *independent increments* property.

Observe that if $N(t)$ is a Poisson process of rate 1, then $N(\lambda t)$ is a Poisson process of rate λ.

Proposition 4. *Let $\{N(J)\}_J$ be a point process that satisfies the independent increments property. Suppose that there is a constant $\lambda > 0$ and a function $f(\varepsilon)$ that converges to 0 as $\varepsilon \to 0$ such that the following holds: For interval J of length ≤ 1,*

$$|P\{N(J)=1\} - \lambda|J|| \leq |J| f(|J|) \quad \textit{and} \tag{7}$$

$$P\{N(J) \geq 2\} \leq |J| f(|J|). \tag{8}$$

Then $\{N(J)\}_J$ is a Poisson *point process with intensity λ.*

Proof. It's only required to prove that the random variable $N(J)$ has a Poisson distribution with mean $\lambda|J|$, because we have assumed independent increments.For this we use — you guessed it — the Law of Small Numbers. Take a large integer n and break J into nonoverlapping subintervals $J_1, J_2, \dots, J_n$ of length $|J|/n$. For each index j define

$$Y_j = \mathbf{1}\{N(J_j) \geq 1\} \quad \text{and}$$
$$Z_j = \mathbf{1}\{N(J_j) \geq 2\}.$$

Clearly,

$$\sum_{j=1}^n Y_j \leq \sum_{j=1}^n N(J_j) = N(J), \quad \text{and} \quad \sum_{j=1}^n Y_j = N(J) \quad \text{if} \quad \sum_{j=1}^n Z_j = 0.$$

Next, I will show that as $n \to \infty$ the probability that $\sum_{j=1}^n Z_j \neq 0$ converges to 0. For this, it is enough to show that the *expectation* of the sum converges to 0 (because for any nonnegative integer-valued random variable W, the probability that $W \geq 1$ is $\leq EW$). But hypothesis (8) implies that $EZ_j \leq (|J|/n) f(|J|/n)$, so

$$E\sum_{j=1}^n Z_j \leq |J| f(|J|/n).$$

Since $f(\varepsilon) \to 0$ as $\varepsilon \to 0$, the claim follows. Consequently, with probability approaching 1 as $n \to \infty$, the random variable $N(J) = \sum_{j=1}^n Y_j$.

Now consider the distribution of the random variable $U_n := \sum_{j=1}^n Y_j$. This is a sum of n independent Bernoulli random variables, all with the same mean EY_j. But hypothesis (8) implies that

$$|EY_j - \lambda|J|/n| \leq (|J|/n) f(|J|/n);$$

consequently, the conditions of the Law of Small Numbers are satisfied, and so the distribution of U_n is, for large n, increasingly close to the Poisson distribution with mean $\lambda|J|$. It follows that $N(J)$ must itself have the Poisson distribution with mean $\lambda|J|$. □

4. INTEROCCURRENCE TIMES OF A POISSON PROCESS

Definition 3. The *occurrence times* $0 = T_0 < T_1 \le T_2 \le \cdots$ of a Poisson process are the successive times that the counting process $N(t)$ jumps, that is, the times t such that $N(t) > N(t-)$. The *interoccurrence times* are the increments $\tau_n := T_n - T_{n-1}$.

Interoccurrence Time Theorem . *(A) The interoccurrence times $\tau_i, \tau_2, \dots$ of a Poisson process with rate λ are independent, identically distributed exponential random variables with mean $1/\lambda$.*

(B) Conversely, let $Y_1, Y_2, \dots$ be independent, identically distributed exponential random variables with mean $1/\lambda$, and define

$$N(t) := \max\{n : \sum_{i=1}^{n} Y_i \le t\}. \tag{9}$$

Then $\{N(t)\}_{t\ge 0}$ is a Poisson process of rate λ.

This is a bit more subtle than some textbooks (notably ROSS) let on. It is easy to see that the *first* occurrence time $T_1 = \tau_1$ in a Poisson process is exponential, because for any $t > 0$

$$P\{T_1 > t\} = P\{N(t) = 0\} = e^{-\lambda t}.$$

It isn't so easy to deduce that the subsequent interoccurrence times are independent exponentials, though, because we don't yet know that Poisson processes "re-start" at their occurrence times. (In fact, this is one of the important consequences of the theorem!) So we'll follow an indirect route that relates *Poisson processes* directly to *Bernoulli processes*. In the usual parlance, a Bernoulli process is just a sequence of i.i.d. Bernoulli random variables (coin tosses); here, however we will want to toss the coin more frequently as the success probability $\to 0$. So let's define a *Bernoulli process* $\{X_r\}_{r\in R}$ *indexed by an arbitrary (denumerable) index set* R to be an assignment of i.i.d Bernoulli random variables to the indices r.

Theorem 5. *(Law of Small Numbers for Bernoulli Processes) For each $m \ge 1$, let $\{X_r^m\}_{r\in\mathbb{N}/m}$ be a Bernoulli process indexed by the integer multiples of $1/m$ with success probability parameter p_m. Let $N^m(t)$ be the corresponding counting process, that is,*

$$N^m(t) = \sum_{r\le t} X_r^m.$$

If $\lim_{m\to\infty} m p_m = \lambda > 0$, then the counting processes $\{N^m(t)\}_{t\ge 0}$ converge in distribution to the counting process of a Poisson process $N(t)$ of rate λ, in the following sense: For any finite set of time points $0 = t_0 < t_1 < \cdots < t_n$,

$$(N^m(t_1), N^m(t_2), \dots, N^m(t_n)) \xrightarrow{\mathscr{D}} (N(t_1), N(t_2), \dots, (N(t_n))) \tag{10}$$

where $\xrightarrow{\mathscr{D}}$ denotes converges in distribution of random vectors.

Proof. The proof is shorter and perhaps easier to comprehend than the statement itself. To prove the convergence (10), it suffices to prove the corresponding result when the random variables $N^m(t_k)$ and $N(t_k)$ are replaced by the *increments* $\Delta_k^m := (N^m(t_k) - N^m(t_{k-1}))$ and $\Delta_k := (N(t_k) - N(t_{k-1}))$. But the convergence in distribution of the increments follows directly from the Law of Small Numbers. □

Proof of the Interoccurrence Time Theorem. The main thing to prove is that the interoccurrence times of a Poisson process are independent, identically distributed exponentials; the second assertion (B) will follow easily from this. By Theorem 5, the Poisson process $N(t)$ is the limit, in the sense (10), of Bernoulli processes $N^m(t)$. It is elementary that *the interoccurrence times of a Bernoulli process are i.i.d. geometric random variables* (scaled by the mesh $1/m$ of the index set). By Proposition 3, these geometric random variables converge to exponentials. Consequently, the interoccurrence times of the limit process $N(t)$ are exponentials.

TECHNICAL NOTE: You may (and should) be wondering how the convergence in distribution of the interoccurrence times follows from the convergence in distribution (10). It's easy, once you realize that the distributions of the *occurrence* times are determined by the values of the random variables $N^m(t)$ and $N(t)$: For instance, the event $T_1^m > t_1$ and $T_2^m > t_2$ is the same as the event $N^m(t_1) < 1$ and $N^m(t_2) < 2$. Thus, once we know (10), it follows that the joint distributions of the occurrence times of the processes $N^m(t)$ converge, as $m \to \infty$, to those of $N(t)$.

This proves (A); it remains to deduce (B). Start with a Poisson process of rate λ, and let τ_i be its interoccurrence times. By what we have just proved, these are i.i.d. exponentials with parameter λ. Now let Y_i be another sequence of i.i.d. exponentials with parameter λ. Since the joint distributions of the sequences τ_i and Y_i are the same, so are the joint distributions of their partial sums, that is,

$$(T_1, T_2, \dots, T_n) \stackrel{\mathcal{D}}{=} (S_1^Y, S_2^Y, \dots, S_n^Y)$$

where

$$S_k^Y = \sum_{j=1}^{k} Y_j.$$

But the joint distribution of the partial sums S_n^Y determines the joint distribution of the occurrence times for the point process defined by (9), just as in the argument above. □

5. POISSON PROCESSES AND THE UNIFORM DISTRIBUTION

Let $\{N(J)\}_J$ be a Poisson point process of intensity $\lambda > 0$. Conditional on the event that $N[0,1] = k$, how are the k points distributed?

Proposition 6. *Given that $N[0,1] = k$, the k points are uniformly distributed on the unit interval $[0,1]$, that is, for any partition $J_1, J_2, \dots, J_m$ of $[0,1]$ into non-overlapping intervals,*

$$P(N(J_i) = k_i \ \forall\, 1 \le i \le m \,|\, N[0,1] = k) = \frac{k!}{k_1! k_2! \cdots k_m!} \prod_{i=1}^{m} |J_i|^{k_i} \tag{11}$$

for any nonnegative integers $k_1, k_2, \dots, k_m$ with sum k.

Proof. The random variables $N(J_i)$ are independent Poisson r.v.s with means $\lambda |J_i|$, by the definition of a Poisson process. Hence, for any nonnegative integers k_i that sum to k,

$$P(N(J_i) = k_i \ \forall\, 1 \le i \le m) = \prod_{i=1}^{m} (\lambda |J_i|)^{k_i} e^{-\lambda |J_i|} / k_i! = \lambda^k e^{-\lambda} \prod_{i=1}^{m} (|J_i|)^{k_i} / k_i!$$

Dividing this by

$$P(N[0,1] = k) = \lambda^k e^{-\lambda} / k!$$

yields equation (11). Finally, to obtain the connection with the uniform distribution, observe that if one were to drop k points independently in $[0,1]$ according to the uniform distribution then the probability that interval J_i would contain exactly k_i points for each $i = 1,2,\dots,m$ would also be given by the right side of equation (11). □

This suggests another way to get a Poisson point process of rate λ: first, construct the counts $N[0,1], N[1,2], N[2,3], \dots$ by i.i.d. sampling from the Poisson distribution with mean λ; then, independently, throw down $N[i,i+1]$ points at random in the interval $[i,i+1]$ according to the uniform distribution. Formally, this construction can be realized on any probability space that supports a sequence $U_1, U_2, \dots$ of independent, identically distributed Uniform-$[0,1]$ random variables and an *independent* sequence $M_1, M_2, \dots$ of independent Poisson-–λ random variables. With these, construct a point process as follows: place the first M_1 points in the interval $[0,1]$ at locations $U_1, U_2, \dots, U_{M_1}$; then place the next M_2 points in $[1,2]$ at locations $1+U_{M_1+1}, \dots, 1+U_{M_1+M_2}$, and so on.

Theorem 7. *The resulting point process is a Poisson point process of rate λ.*

Proof. What must be shown is that the point process built this way has the independent increments property, and that the increments $N(t+s)-N(t)$ have Poisson distributions with means λs. Without loss of generality, we need only consider increments $N(t+s)-N(t)$ for which the endpoints t and $t+s$ lie in the same interval $[n,n+1]$ (because if $[t,t+s]$ straddles n, it can be broken into $[t,n]\cup[n,t+s]$). For simplicity, let's restrict our attention to intervals that lie in $[0,1]$. Suppose, then, that

$$0 = t_0 < t_1 < \cdots < t_n = 1.$$

Let Δ_k = the number of uniforms U_i that end up in the interval $J_k := (t_{k-1}, t_k]$. Then Δ_k is the number of "successes" in M_1 Bernoulli trials, where a trial U_i is considered a success if $U_i \in J_k$ and a failure otherwise. The success probability is the length of the interval J_k. Hence, the Thinning Law implies that Δ_k has the Poisson distribution with mean $\lambda|J_k|$. Similarly, the Generalized Thinning Law implies that the increments Δ_k are mutually independent. □

Corollary 8. *Let $T_1, T_2, \dots$ be the occurrence times in a Poisson process $N(t)$ of rate λ. Then conditional on the event $N(1) = m$, the random variables $T_1, \dots, T_m$ are distributed in the same manner as the order statistics of a sample of m i.i.d. uniform-$[0,1]$ random variables.*

6. Poisson Point Processes

Definition 4. A *point process* on $\mathbb{R}^k$ is a mapping $J \mapsto N_J = N(J)$ that assigns to each reasonable (i.e., Borel) subset $J \subset \mathbb{R}^k$ a nonnegative, integer-valued random variable N_J in such a way that if $J_1, J_2, ..$ are pairwise disjoint then

$$N(\cup_i J_i) = \sum_i N(J_i) \tag{12}$$

A *Poisson point process* on $\mathbb{R}^k$ with intensity function $\lambda : \mathbb{R}^k \to [0,\infty)$ is a point process N_J that satisfies the following two additional properties:

(A) If J is such that $\Lambda(J) := \int_J \lambda(x)\,dx < \infty$ then N_J has the Poisson distribution wwith mean $\Lambda(J)$.
(B) If $J_1, J_2, \dots$ are pairwise disjoint then the random variables $N(J_i)$ are mutually independent.

The intensity function $\lambda(x)$ need not be integrable or continuous, but in most situations it will be *locally integrable*, that is, its integral over any bounded rectangle is finite. The definition generalizes easily to allow intensities that are *measures* Λ rather than *functions* λ.

Proposition 9. *To prove that a point process is a Poisson point process, it suffices to verify conditions (A)–(B) for* rectangles J, J_i *with sides parallel to the coordinate axes.*

I won't prove this, as it would require some measure theory to do so. The result is quite useful, though, because in many situations rectangles are easier to deal with than arbitrary Borel sets.

Example: The $M/G/\infty$ Queue. Jobs arrive at a service station at the times of a Poisson process (on the timeline $[0,\infty)$) with rate $\lambda > 0$. To each job is attached a *service time*; the service times Y_i are independent, identically distributed with common density f. The service station has infinitely many servers, so that work on each incoming job begins immediately. Once a job's service time is up, it leaves the system. We would like to know, among other things, the distribution of the number Z_t of jobs in the system at time t.

The trick is to realize that the service times Y_i and the arrival times A_i determine a *two-dimensional* Poisson point process N_J with intensity function $\lambda(t,y) = \lambda f(y)$. The random variable N_J is just the number of points (A_i, Y_i) that fall in J. To see that this is in fact a Poisson point process, we use Proposition 9. Suppose that $J = [s,t]\times[a,b]$ is a rectangle; then N_J is the number of jobs that arrive during the time interval $[s,t]$ whose service times are between a and b. Now the total number of jobs that arrive during $[s,t]$ is Poisson with mean $\lambda(t-s)$, because by hypothesis the arrival process is Poisson. Each of these tosses a coin with success probability $p = \int_a^b f(y)\,dy$ to determine whether its service time is between a and b. Consequently, the Thinning Law implies that N_J has the right Poisson distribution. A similar argument show that the "independent increments" property (B) holds for rectangles.

Let's consider the random variable Z_t that counts the number of jobs in the system at time t. If a job arrives at time r and its service time is s, then it will be in the system from time r until time $r+s$. This can be visualized by drawing a line segment of slope -1 starting at (r,s) and extending to $(r+s,0)$: the interval of the real line lying underneath this line segment is $[r,r+s]$. Thus, the number of jobs in the system at time t can be gotten by counting the number of points (A_i, Y_i) with $A_i \le t$ that lie *above* the line of slope -1 through $(t,0)$. (See the accompanying figure.) It follows that Z_t has the Poisson distribution with mean

$$\int_{0,t}\int_{t-s}^{\infty}\lambda f(y)\,dy.$$

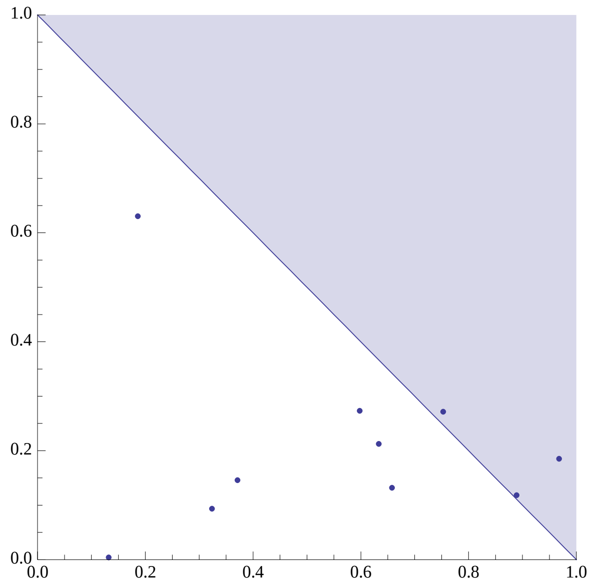

FIGURE 1. $M/G/\infty$ Queue