Some notes on the Poisson distribution

Ernie Croot

October 7, 2010

1   Introduction

The Poisson distribution is one of the most important that we will encounter in this course – it is right up there with the normal distribution. Yet, because of time limitations, and due to the fact that its true applications are quite technical for a course such as ours, we will not have the time to discuss them. The purpose of this note is not to address these applications, but rather to provide source material for some things I presented in my lectures.

2   Basic properties

Recall that X is a Poisson random variable with parameter λ if it takes on the values $0, 1, 2, ...$ according to the probability distribution

$$p(x) \; = \; P(X = x) \; = \; \frac{e^{-\lambda}\lambda^x}{x!}.$$

By convention, $0! = 1$.

Let us verify that this is indeed a legal probability density function (or "mass function" as your book likes to say) by showing that the sum of $p(n)$ over all $n \geq 0$, is 1. We have

$$\sum_{n=0}^{\infty} p(n) \; = \; e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!}.$$

This last sum clearly is the power series formula for e^λ, so

$$\sum_{n=0}^{\infty} p(n) \; = \; e^{-\lambda}e^{\lambda} \; = \; 1,$$

as claimed.

A basic fact about the Poisson random variable X (actually, two facts in one) is as follows:

Fact. $E(X) = V(X) = \lambda$.

Let us just prove here that $E(X) = \lambda$: As with showing that $p(x)$ is a legal pdf, this is a simple exercise in series manipulation. We begin by noting that

$$E(X) = \sum_{n=0}^{\infty} np(n) = \sum_{n=1}^{\infty} n \cdot \frac{e^{-\lambda}\lambda^n}{n!}.$$

Noting that

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 3 \cdot 2 \cdot 1,$$

we see that we may cancel the n with the $n!$ to leave an $(n-1)!$ in the denominator. So, pulling out the $e^{-\lambda}$ and cancelling the n, we deduce that

$$E(X) = e^{-\lambda} \sum_{n=1}^{\infty} \frac{\lambda^n}{(n-1)!} = \lambda e^{-\lambda} \sum_{n=1}^{\infty} \frac{\lambda^{n-1}}{(n-1)!}.$$

(Note here that the term $n = 1$ has a 0! in the denominator.)

Upon setting $m = n - 1$, we find that

$$E(X) = \lambda e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = \lambda e^{-\lambda} e^{\lambda} = \lambda,$$

and we are done.

3 A sum property of Poisson random variables

Here we will show that if Y and Z are independent Poisson random variables with parameters λ_1 and λ_2, respectively, then $Y + Z$ has a Poisson distribution with parameter $\lambda_1 + \lambda_2$.

Before we even begin showing this, let us recall what it means for two discrete random variables to be "independent": We say that Y and Z are independent if

$$P(Y = y, Z = z) \;=\; P(Y = y \text{ and } Z = z) \;=\; P(Y = y)P(Z = z).$$

This is not really the "official" definition of independent random variables, but it is good enough for us.

Letting now
$$X \;=\; Y + Z,$$
we have that

$$
\begin{aligned}
P(X = x) \;=\; P(Y + Z = x) \;&=\; \sum_{\substack{y,z \geq 0 \\ y+z=x}} P(Y = y, Z = z) \\
&=\; \sum_{\substack{y,z \geq 0 \\ y+z=x}} P(Y = y)P(Z = z) \\
&=\; \sum_{\substack{y,z \geq 0 \\ y+z=x}} \frac{e^{-\lambda_1}\lambda_1^y}{y!} \cdot \frac{e^{-\lambda_2}\lambda_2^z}{z!} \\
&=\; e^{-(\lambda_1+\lambda_2)} \sum_{\substack{y,z \geq 0 \\ y+z=x}} \frac{\lambda_1^y \lambda_2^z}{y!z!}.
\end{aligned}
$$

Now we require a little trick to finish our proof. We begin by recognizing that the $y!z!$ is the denominator of

$$\frac{x!}{y!z!} \;=\; \binom{x}{y} \;=\; \binom{x}{z}.$$

So, if we insert a factor $x!$ in the sum, then we will have some binomial coefficient. Indeed, we have

$$P(X = x) \;=\; \frac{e^{-(\lambda_1+\lambda_2)}}{x!} \sum_{\substack{y,z \geq 0 \\ y+z=x}} \binom{x}{y} \lambda_1^y \lambda_2^z.$$

Notice that from the binomial theorem we deduce that this last sum is just

$$(\lambda_1 + \lambda_2)^x.$$

So,
$$P(X = x) \;=\; \frac{e^{-(\lambda_1 + \lambda_2)}(\lambda_1 + \lambda_2)^x}{x!},$$
and therefore X has a Poisson distribution with parameter $\lambda_1 + \lambda_2$, as claimed.

4 An application

It turns out that prime numbers in short intervals obey a "Poisson distribution". To discuss this, I need to tell you about the Prime Number Theorem:

Prime Number Theorem. The number of primes $p \leq x$ has size about $x/\log x$, where here the log is to the base-e. That is, if $\pi(x)$ denotes this number of primes, then

$$\pi(x) \;\sim\; x/\log x, \text{ which means that } \lim_{x \to \infty} \frac{\pi(x)}{x/\log x} \;=\; 1.$$

In what follows, we adopt the notation $\pi(I)$ to denote the number of primes in a given interval I.

Now suppose that x is a "really large" number, say $x = 10^{100}$. Then, if you pick a number $n \leq x$ at random, there is about a $1/100 \log(10)$ (about 0.43 percent) chance that it will be prime; and we expect that a typical interval $[n, n + \log n] \subset [1, x]$ will contain about one prime.

In fact, much more is true: If you pick $n \leq x$ at random, and choose $\lambda > 0$ and j not too large (say $\lambda, j \leq 20$), then the number of primes in $[n, n + \lambda \log n]$ roughly obeys a Poisson distribution:

$$\mathbb{P}(\pi([n, n + \lambda \log n]) = j) \;\approx\; \frac{e^{-\lambda}\lambda^j}{j!}.$$

Notice that I didn't use equality here; in order to get equality we would have to let $x \to \infty$ in some way – certainly the larger we take x to be, the closer the above probability comes to $e^{-\lambda}\lambda^j/j!$.

Given what you know about Poisson distributions, it is an interesting exercise to figure out *why* one would expect the primes to be "Poisson-distributed" as stated above.