The Poisson Distribution

Bret Larget
Departments of Botany and of Statistics
University of Wisconsin—Madison
Statistics 371
23rd September 2005

The Poisson Distribution

The Poisson distribution arises in many biological contexts. Examples of random variables for which a Poisson distribution might be reasonable include:

- the number of bacterial colonies in a Petri dish;
- the number of trees in an area of land;
- the number of offspring an individual has;
- the number of nucleotide base substitutions in a gene over a period of time;

Probability Mass Function

The probability mass function of the Poisson distribution with mean $\mu$ is

$$
Pr\{Y = k | \mu\} = \frac{e^{-\mu} \mu^k}{k!} \quad \text{for } k = 0, 1, 2, \ldots.
$$

The Poisson distribution is discrete, like the binomial distribution, but has only a single parameter $\mu$ that is both the mean and the variance.

Example

Suppose the number of individual plants of a given species we expect to find in a one meter square quadrat follows the Poisson distribution with mean $\mu = 10$. Find the probability of finding exactly 12 individuals.

$Y \sim \text{Poisson}(10)$.

$$
Pr\{Y = 12 | \mu = 10\} = \frac{e^{-10} \cdot 10^{12}}{12!} \approx 0.0948.
$$
Example in R

- In R, you can compute Poisson probabilities with the function `dpois`.
- For the previous example, try the following.

```r
> dpois(12, 10)
[1] 0.09478033
```

Poisson approximation to the Binomial

- The Poisson distribution is a good approximation to the binomial distribution when $p$ is small.
- The approximation is better for large $n$.
- If $p$ is small, then the binomial probability of exactly $k$ successes is approximately the same as the Poisson probability of $k$ with $\mu = np$.

Example

Suppose that $Y \sim \text{Binomial}(1000, 0.01)$. Find $\Pr\{Y = 8\}$.
The exact calculation is:

$$
\Pr\{Y = 8\} = \frac{1000!}{8!992!}(0.01)^8(0.99)^{992} \approx 0.112824
$$

Working with large factorials can be messy. The Poisson approximation uses $\mu = 1000 \times 0.01 = 10$ and is:

$$
\Pr\{Y = 8\} \approx \frac{e^{-10} \cdot 10^8}{8!} \approx 0.112599
$$

Example using R

Here is the same example using R.

```r
> dbinom(8, 1000, 0.01)
[1] 0.1128241
> dpois(8, 1000 * 0.01)
[1] 0.1125990
```
The Poisson Process

- The Poisson Process arises naturally under assumptions that are often reasonable.
- For the following, think of points as being exact times or locations.
- The assumptions are:
  - The chance of two simultaneous points is negligible;
  - The expected value of the random number of points in a region is proportional to the size of the region.
  - The random number of points in non-overlapping regions are independent.
- Under these assumptions, the random variable that counts the number of points has a Poisson distribution.
- If the expected rate of points is \( \lambda \) points per unit length (area), then the distribution of the number of points in an interval (region) of size \( t \) is \( \mu = \lambda t \).

Example

Suppose that we assume that at a location, a particular species of plant is distributed according to a Poisson process with expected density 0.2 individuals per square meter. In a nine square meter quadrat, what is the probability of no individuals?

**Solution:** The number of individuals has a Poisson distribution with mean \( \mu = 9 \times 0.2 = 1.8 \). The probability of this is

\[
Pr\{Y = 0 \mid \mu = 1.8\} = \frac{e^{-1.8}(1.8)^0}{0!} = 0.165299
\]

In R, we can compute this as

```r
> dpois(0, 1.8)
[1] 0.1652989
```

Example (cont.)

Find the probability of three or more individuals.

**Solution:** Instead of summing the probabilities from 3 to infinity, we can use the complement rule.

\[
Pr\{Y \geq 3\} = 1 - Pr\{Y \leq 2\} = 1 - Pr\{Y = 0\} - Pr\{Y = 1\} - Pr\{Y = 2\}
\]

In R, this is found by one of two ways.

```r
> 1 - ppois(2, 1.8)
[1] 0.2693789
> 1 - sum(dpois(0:2, 1.8))
[1] 0.2693789
```