13 POISSON DISTRIBUTION

Examples

- You have observed that the number of hits to your web site occur at a rate of 2 a day.
 Let X be be the number of hits in a day
- You observe that the number of telephone calls that arrive each day on your mobile phone over a period of a year, and note that the average is 3.
 Let X be the number of calls that arrive in any one day.
- Records show that the average rate of job submissions in a busy computer centre is 4 per minute. Let X be the number of jobs arriving in any one minute.
- 4. Records indicate that messages arrive to a computer server at the rate of 6 per hour.Let X be the number of messages arriving in any one hour.

Generally

X = number of events, distributed independently in time, occurring in a fixed time interval.

X is a Poisson variable with pdf:

$$P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}, \quad x = 0, 1, \dots, \infty$$

where λ is the average.

Example:

Consider a computer system with Poisson job-arrival stream at an average of 2 per minute. Determine the probability that in any one-minute interval there will be

(i) 0 jobs;

- (ii) exactly 2 jobs;
- (iii) at most 3 arrivals.

(iv) What is the maximum jobs that should arrive one minute with 90 % certainty?

Solution: Job Arrivals with $\lambda = 2$ (i) No job arrivals:

$$P(X=0) = e^{-2} = .135$$

In R

dpois(0, 2)
[1] 0.1353353

(ii) Exactly 3 job arrivals:

$$P(X=3) = e^{-2}\frac{2^3}{3!} = .18$$

In R

dpois(3, 2)
[1] 0.1804470

(iii) At most 3 arrivals

$$P(X \le 3) = P(0) + P(1) + P(2) + P(3)$$

= $e^{-2} + e^{-2}\frac{2}{1} + e^{-2}\frac{2^2}{2!} + e^{-2}\frac{2^3}{3!}$
= $0.1353 + 0.2707 + 0.2707 + 0.1805$
= 0.8571

 $\ln R$

ppois(3,2)
[1] 0.8571235

more than 3 arrivals:

$$P(X > 3) = 1 - P(X \le 3)$$

= 1 - 0.8571
= 0.1429

(iv) Maximum arrivals with at least 90% certainty:

i.e. 90% quantile

Choose k so that

$$P(X \le k) \ge .9$$

In *R*

qpois(.9, 2)
[1] 4

at least a 90% chance that the number of job submissions in any minute does not exceed 4.

equivalently

less than a 10% chance that there will be more than 4 job submissions in any one minute.





Calls to Mobile: Poisson(3)

Job Submissions: Poisson(4) (x = X) (x = 0.0) (x = 0.0)(x = 0.0)

Number of Submissions

6

8

10

12

0

2

4

Messages to Server: Poisson(6)



Poisson Probability Density Functions

```
par(mfrow = c(2,2)) # multiframe
x<-0:12 #look at the first 12 probabilities
plot (x, dpois(x, 2),
    xlab = "Number of Hits", ylab = "P(X = x)",
    type = "h", main= "Web Site Hits: Poisson(2)")
plot (x, dpois(x, 3),
    xlab = "Number of Calls", ylab = "P(X = x)",
    type = "h", main= "Calls to Mobile: Poisson(3)")
plot (x, dpois(x, 4),
    xlab = "Number of Submissions", ylab = "P(X = x)",
    type = "h", main= "Job Submissions: Poisson(4)")
plot (x, dpois(x, 6),
    xlab = "Number of Messages", ylab = "P(X = x)",
    type = "h", main= "Messages to Server: Poisson(6)")
```

Poisson Cumulative Distribution Functions









Number of Messages

Poisson Cumulative Distribution Functions

```
par(mfrow = c(2,2)) # multiframe
x<-0:12
plot (x, ppois(x, 2),
    xlab = "Number of Hits", ylab = "P(X = x)",
    type = "s", main= "Web Site Hits:lambda=2")
plot (x, ppois(x, 3),
    xlab = "Number of Calls", ylab = "P(X = x)",
    type = "s", main= "Calls to Mobile:lambda=3")
plot (x, ppois(x, 4),
    xlab = "Number of Submissions", ylab = "P(X = x)",
    type = "s", main= "Submissions:lambda=4")
plot (x, ppois(x, 6),
    xlab = "Number of Messages", ylab = "P(X = x)",
    type = "s", main= "Server Messages:lambda=6", )
```

Derivations of Some Properties of Poisson

1. Clearly

$$e^{-\lambda} \frac{\lambda^x}{x!} > 0$$
 since $\lambda > 0$

Also

$$\sum_{x=0}^{\infty} e^{-\lambda} \frac{\lambda^x}{x!} = 1$$

since

$$e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots$$

i.e.

$$\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{\lambda}$$

2. $E(X) = \lambda$

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

= $e^{-\lambda} \sum_{x=1}^{\infty} x \frac{\lambda^x}{x!}$
= $e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!}$
= $e^{-\lambda} \lambda \left[\sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \right]$
= $e^{-\lambda} \lambda \left[1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots \right]$
= $e^{-\lambda} \lambda e^{\lambda} = \lambda$

APPLICATIONS OF THE POISSON

The Poisson distribution arises in two ways:

1. Events distributed independently of one another in time:

 $\mathsf{X}=\mathsf{the}$ number of events occurring in a fixed time interval has a Poisson distribution.

$$PDF: \quad p(x) = e^{-\lambda} \frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \dots; \lambda > 0$$

Example: X = the number of telephone calls in an hour.

2. As an approximation to the binomial when p is small and n is large,

When examining the number of defectives in a large batch where p, the defective rate, is usually small.

The manufacturer of the disk drives in one of the well-known brands of microcomputers expects 2% of the disk drives to malfunction during the microccomputer's warranty period.

Calculate the probability that in a sample of 100 disk drives, that not more than three will malfunction.

No. of disk drives	Binomial	Poisson
malfunctioning		Approximation
k	$\binom{100}{k}.02^k.98^{100-k}$	$e^{-2}2^k/k!$
0	0.13262	0.13534
1	0.27065	0.27067
2	0.27341	0.27067
3	0.18228	0.18045
Total	0.85890	0.85713

Poisson as an approximation to the binomial when n is large p is small Recall:

- mean of binomial = np
- mean of Poisson = λ

PDF of Binomial

$$P(x) = \binom{n}{x} p^{x} (1-p)^{n-x}; \quad p = \frac{\lambda}{n}$$
$$= \binom{n}{x} \left(\frac{\lambda}{n}\right)^{x} \left(1-\frac{\lambda}{n}\right)^{n-x}$$
$$= \frac{n!}{x!(n-x)!} \frac{\lambda^{x}}{n^{x}} \frac{(1-\frac{\lambda}{n})^{n}}{(1-\frac{\lambda}{n})^{x}} \to \frac{\lambda^{x}}{x!} e^{-\lambda}$$

Now

$$\begin{aligned} (1-\frac{\lambda}{n})^n &\to e^{-\lambda} \text{ as } n \to \infty\\ (1-\frac{\lambda}{n})^x &\to 1 \text{ as } n \to \infty\\ \\ \frac{n!}{(n-x)!n^x} &= \frac{n(n-1)\dots(n-(x-1))}{n^x}\\ &= 1(1-\frac{1}{n})(1-\frac{2}{n})\dots(1-\frac{x-1}{n}) \to 1 \end{aligned}$$

Some Examples

1. Messages arrive to a server at the rate of 6 per hour. What is the maximum number k so that the probability that the number of messages to the server in an hour more than this value is .75

Solution

```
qpois(.75, 6)
[1] 8
At least a 75% chance of \leq 8 messages
25% change of > 8; i.e. 3rd quartile.
For 1st quartile,
qpois(.25, 6)
[1] 4
Interquartile range: [4, 8],
```

i.e.

50% of the time, arrivals will be between 4 and 8.

A web server receives an average of 1000 queries per day. How much excess capacity should it build in to be 95% sure that it can meet the demand?

Poisson with $\lambda = 1000$

Choose k so that

 $P(X \le k) \ge .95$

qpois(.95, 1000)
[1] 1052

and to double check

ppois(1052, 1000)
[1] 0.9506515

The extra capacity needed is 52.

The mean number of errors due to a particular bug occurring in a minute is 0.0001

- 1. What is the probability that no error will occur in 20 minutes?
- 2. How long would the program need to run to ensure that there will be a 99.95% chance that an error will show up to highlight this bug?

Solution

Poisson $\lambda = .0001$ per minute $\lambda = 20(.0001) = .002$ per 20 minute interval.

probability that no error will occur in 20 min:

$$P(0) = e^{-.002} = 0.998002$$

99.88% chance OF 0 in 20 min.

Equivalently

1.12% that an error will show up in the first 20 minutes.

For 99.95% sure of catching bug:

$$P(X \ge 1) \ge .9995,$$

or equivalently

$$P(X = 0) = .0005$$

 $P(No \text{ occurrence in } k \text{ minutes}) = e^{-(.0001)k}$

To be 99.95% sure, choose k so that

 $e^{-(.0001)k} \le .0005$

or equivalently

$$1 - e^{-(.0001)k} > .9995$$

k > 75,000 mins k ; 75000/60 = 1250 hours k . 1250/24 = 52 days

In R we can examine the probabilities of at least one error in differing timespans:

```
k<-seq(50000,100000, 5000) #Running time in 5,000 int.
y<-1-exp(-(.0001)*k) #At least one in k minutes
plot(k,y, xlab = "Running time of the package in minutes",
   ylab = "Probability of at least one bug")
abline(h = .9995)
```

Bug Detection.



The average number of defects per wafer (defect density) is 3. The redundancy built into the design allows for up to 4 defects per wafer. What is the probability that the redundancy will not be sufficient if the defects follow a Poisson distribution?

Poisson with $\lambda = 4$.

For k defects:

$$P(X=k) = e^{-3}\frac{3^k}{k}$$

where \boldsymbol{X} be the number of defects per wafer.

The redundancy will not be sufficient when X > 4.

$$\begin{split} &P(X>4) \\ = \ 1 - (P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4)) \\ = \ 1 - e^{-3}(1 + \frac{3^1}{1} + \frac{3^2}{2!} + \frac{3^3}{3!} + \frac{3^4}{4!}) \end{split}$$

In *R* we obtain P(X > 4) using ppois

1-ppois(4, 3) [1] 0.1847368

over an 18% chance that the defects will exceed redundancy.

${\it R}$ Functions for the Poisson Distribution

```
• dpois

dpois(x = 4, lambda = 3)

or

dpois(4, 3)

P(X = 4) \lambda = 3
```

• ppois

ppois(x = 4, \lambda = 3) or ppois(4, 3) $P(X \le 4)$ with $\lambda = 3$

• qpois

qpois(.95, lambda = 3) or qpois(.95, 3) [1] 7 Choose k so that $P(X \le k) \ge .95.$

> qpois(.95, 3) [1] 6