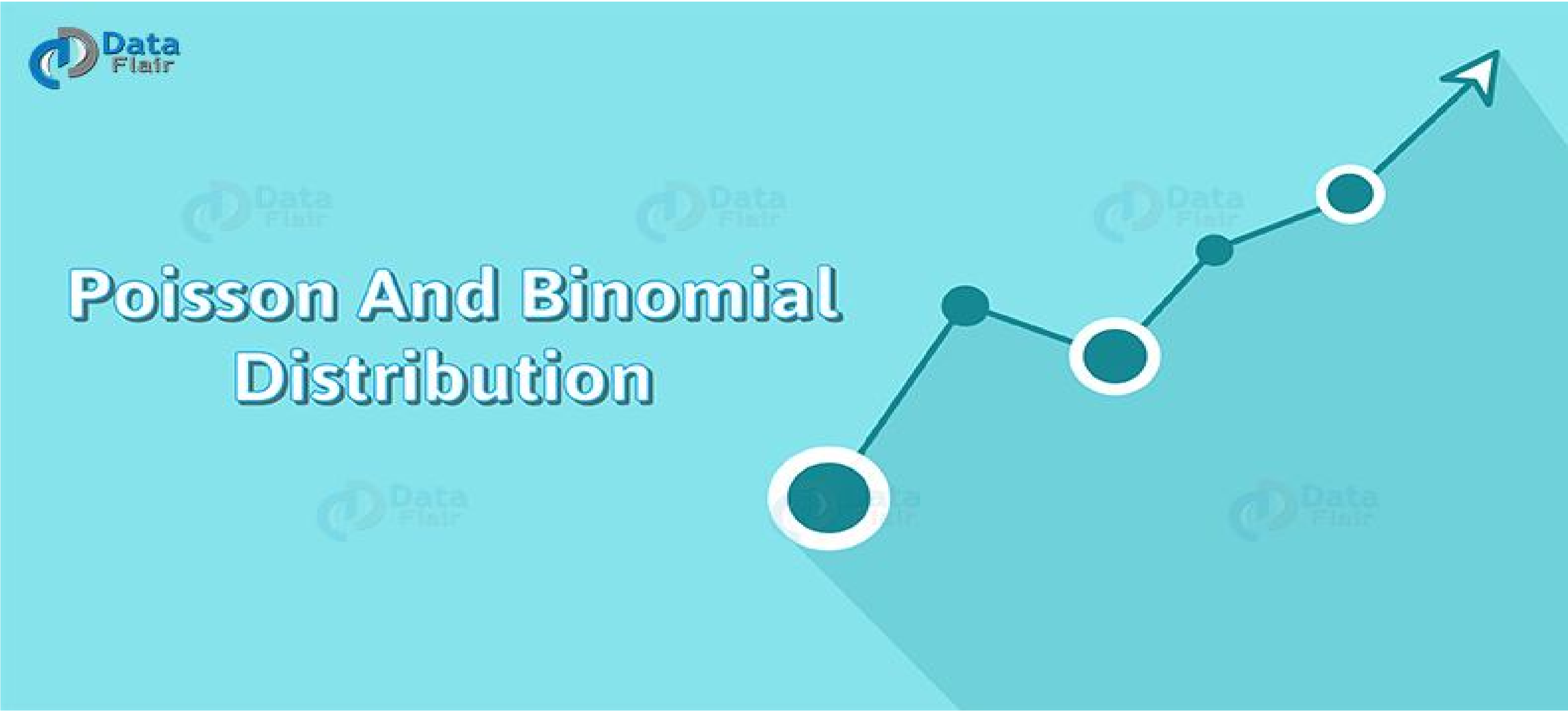
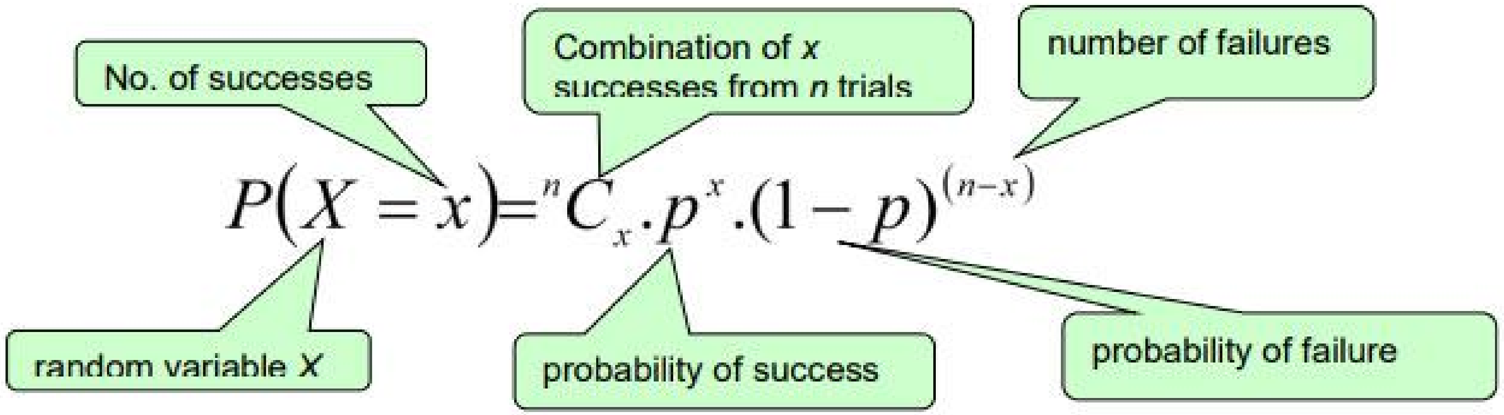

Poisson And Binomial
Distribution

# Probability Distribution

A probability distribution describes how the values of a random variable is distributed. For example, the collection of all possible outcomes of a sequence of coin tossing is known to follow the Binomial distribution AND Poisson distribution. Whereas the means of sufficiently large samples of a data population are known to resemble the normal distribution . Since the characteristics of these theoretical distributions are well understood, they can be used to make statistical inferences on the entire data population as a whole.

# Binomial Distribution

- The binomial distribution is a discrete probability distribution. It describes the outcome of n independent trials in an experiment. Each trial is assumed to have only two outcomes, either success or failure. If the probability of a successful trial is p, then the probability of having x successful outcomes in an experiment of n independent trials is as follows.



# R-Syntax

R has four in-built functions to generate binomial distribution. They are described below.

```
dbinom(x, size, prob)
pbinom(x, size, prob)
qbinom(p, size, prob)
rbinom(n, size, prob)
```

- **x** is a vector of numbers.
- **p** is a vector of probabilities.
- **n** is number of observations.
- **size** is the number of trials.
- **prob** is the probability of success of each trial.

- **dbinom()** This function gives the probability density distribution at each point.
- Both of the R commands in the box below do exactly the same thing.

```
> dbinom(27, size=100, prob=0.25)
[1] 0.08064075
> dbinom(27, 100, 0.25)
[1] 0.08064075
> dbinom(4,size=12,prob=1/6)
[1] 0.08882807
> dbinom(4,12,1/6)
[1] 0.08882807
```

- **pbinom()** This function gives the cumulative probability of an event. It is a single value representing the probability.
- *What is P(X <= 1) when X has the Bin(25, 0.005) distribution?*
- *What is P(X <= 27) when X is has the Bin(100, 0.25) distribution?*

```
> pbinom(1, 25, 0.005)
[1] 0.9930519
> pbinom(27, 100, 0.25)
[1] 0.7223805
```

- **qbinom()** This function takes the probability value and gives a number whose cumulative value matches the probability value.
- What are the 10th, 20th, and so forth quantiles of the Bin(10, 1/3) distribution?
- How many heads will have a probability of 0.25 will come out when a coin is tossed 51 times.

```
> qbinom(0.1, 10, 1/3)
[1] 1
> qbinom(0.2, 10, 1/3)
[1] 2
> qbinom(seq(0.1, 0.9, 0.1), 10, 1/3)
[1] 1 2 3 3 3 4 4 5 5
> x <- qbinom(0.25,51,1/2)
> x
[1] 23
```

- **rbinom()** This function generates required number of random values of given probability from a given sample.

```
> # Find 8 random values from a sample of 150 with probability of 0.4.
> x <- rbinom(8,150,.4)
> x
[1] 71 60 61 72 68 65 58 49
```

- **Problem1:**

- If a committee has 7 members, find the probability of having more female members than male members given that the probability of having a male or a female member is equal.

Sol: The probability of having a female member = 0.5

The probability of having a male member = 0.5

To have more female members, the number of females should be greater than or equal to 4.

```
> 1-pbinom(3,7,0.5)
[1] 0.5
```

## › Problem 1:

1. Suppose X ~ Bin(10, 0.4), what is P(X = 7)?

$$\begin{aligned} P(X=7) &= {}^{10}C_7(0.4)^7(1-0.4)^{(10-7)} \\ &= (120)(0.4)^7(0.6)^3 \\ &= 0.0425 \end{aligned}$$

```
> dbinom(7,10,0.4)
[1] 0.04246733
```

## › Problem 2:

2. Suppose Y ~ Bin(8, 0.15), what is P(Y < 3)?

$$\begin{aligned} P(Y<3) &= P(Y=0)+P(Y=1)+P(Y=2) \\ &= {}^{8}C_0(0.15)^0(0.85)^8 + {}^{8}C_1(0.15)^1(0.85)^7 + {}^{8}C_2(0.15)^2(0.85)^6 \\ &= 0.2725+0.3847+0.2376 \\ &= 0.8948 \end{aligned}$$

```
> pbinom(2,8,0.15)
[1] 0.8947872
```

› Problem 3:

3. Suppose W ~ Bin(50, 0.12), what is P(W > 2)?

$$\begin{aligned}
P(W > 2) &= P(W = 3) + P(W = 4) + \ldots + P(W = 50) \\
&= 1 - P(W \leq 2) \\
&= 1 - \Big(P(W = 0) + P(W = 1) + P(W = 2)\Big) \\
&= 1 - \Big({}^{50}C_0(0.12)^0(0.88)^{50} + {}^{50}C_1(0.12)^1(0.88)^{49} + {}^{50}C_2(0.12)^2(0.88)^{48}\Big) \\
&= 1 - \Big(0.00168 + 0.01142 + 0.03817\Big) \\
&= 0.94874
\end{aligned}$$

```
> 1-pbinom(2,50,0.12)
[1] 0.9487358
```

## › Problem 4:

› In a box of switches it is known 10% of the switches are faulty. A technician is wiring 30 circuits, each of which needs one switch. What is the probability that (a) all 30 work, (b) at most 2 of the circuits do not work?

(a) Probability that all 30 work is $\mathsf{P}(X = 30) = {}^{30}C_{30}(0.9)^{30}(0.1)^{0} = 0.04239$

(b) The statement that "at most 2 circuits do not work" implies that 28, 29 or 30 work. That is $X \geq 28$

$$\begin{aligned}
\mathsf{P}(X \geq 28) &= \mathsf{P}(X = 28) + \mathsf{P}(X = 29) + \mathsf{P}(X = 30) \\
\mathsf{P}(X = 30) &= {}^{30}C_{30}(0.9)^{30}(0.1)^{0} = 0.04239 \\
\mathsf{P}(X = 29) &= {}^{30}C_{29}(0.9)^{29}(0.1)^{1} = 0.14130 \\
\mathsf{P}(X = 28) &= {}^{30}C_{28}(0.9)^{28}(0.1)^{2} = 0.22766
\end{aligned}$$

Hence $\mathsf{P}(X \geq 28) = 0.41135$

```
> dbinom(30,30,0.9)
[1] 0.04239116
```

```
> 1-pbinom(27,30,0.9)
[1] 0.4113512
```

- Problem 5:
- If 10% of the Screws produced by an automatic machine are defective, find the probability that out of 20 screws selected at random, there are
- (i) Exactly 2 defective (ii) At least 2 defectives
- (iii) Between 1 and 3 defectives (inclusive)

```
> # Exactly 2 defective
> dbinom(2,20,0.10)
[1] 0.2851798
> #At least two defectives
> 1-dbinom(1,20,0.10)
[1] 0.7298297
> #Between 1 and 3 defectives (inclusive)
> sum(dbinom(1:3,20,0.10))
[1] 0.74547
>
```

# Poisson distribution

In probability theory, the Poisson distribution is a very common discrete probability distribution. A Poisson distribution helps in describing the chances of occurrence of a number of events in some given time interval or given space conditionally that the value of average number of occurrence of the event is known. This is a major and only condition of Poisson distribution.

The random variable X is said to follow the Poisson distribution if and only if

$$p[X = x] = \frac{e^{-\lambda}\lambda^{x}}{\lfloor x}, x = 0, 1, 2, .....$$

Conditions :-

(i) Number of Bernoulli trials (n) is indefinitely large, ($n \to \infty$)
(ii) The trials are independent.
(iii) Probability of success (p) is very small, ($p \to 0$)

# R syntax

- dpois(x, lambda, log = FALSE)
- ppois(q, lambda, lower.tail = TRUE, log.p = FALSE)
- qpois(p, lambda, lower.tail = TRUE, log.p = FALSE)
- pois(n, lambda)

› Practice problems:

1. What is P(Y = 4)?

```
> dpois(4, lambda = 2.6)
[1] 0.1414218
```

2. What is P(Y >= 2)?

```
> 1 - ppois(1, lambda = 2.6)
[1] 0.7326151
```

3. What is P(3 <= Y <= 6)?

```
> ppois(6, lambda = 2.6) - ppois(2, lambda = 2.6)
[1] 0.4644003
```

4. 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.

Solution: Job Arrivals with $\lambda = 2$

(i) No job arrivals:

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

```
> dpois(0, 2)
[1] 0.1353353
```

(ii) Exactly 3 job arrivals:

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

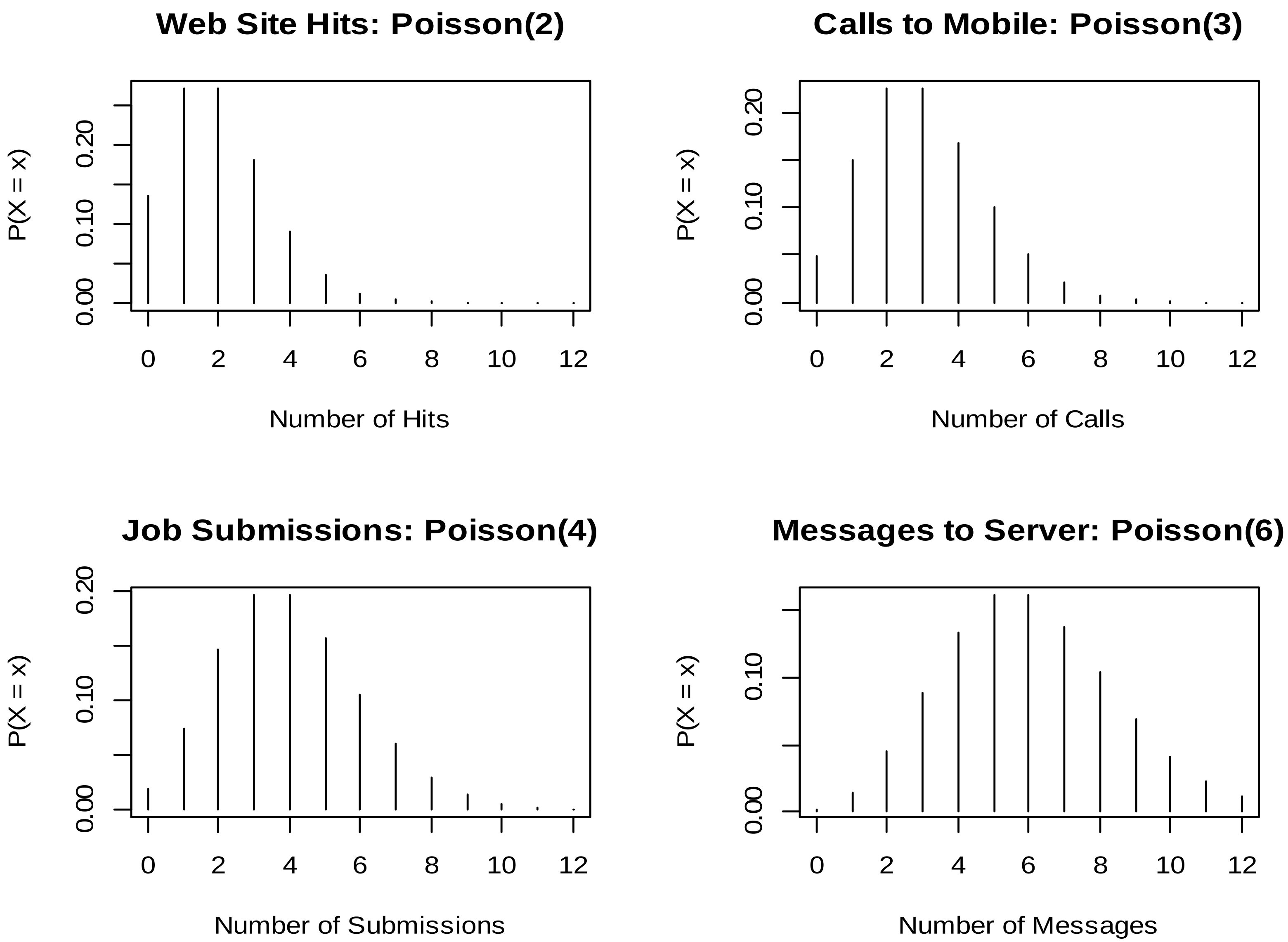
```
> dpois(3, 2)
[1] 0.180447
```

(iii) At most 3 arrivals

$$\begin{aligned} P(X \leq 3) &= P(0) + P(1) + P(2) + P(3) \\ &= e^{-2} + e^{-2}\frac{2}{1} + e^{-2}\frac{2^2}{2!} + e^{-2}\frac{2^3}{3!} \\ &= 0.1353 + 0.2707 + 0.2707 + 0.1805 \\ &= 0.8571 \end{aligned}$$

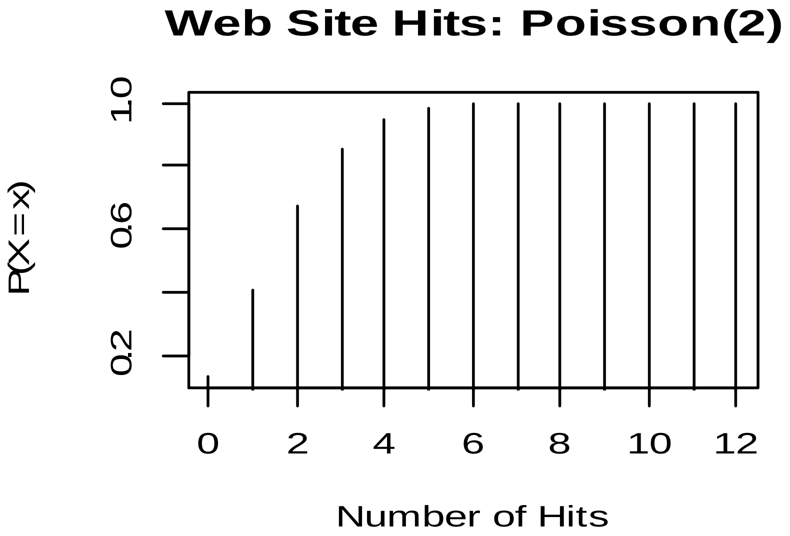
```
> ppois(3,2)
[1] 0.8571235
```

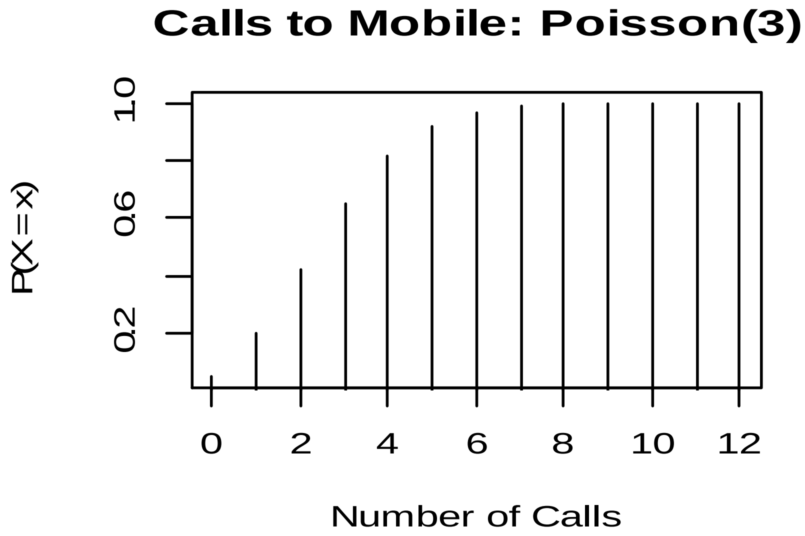
- Poisson Probability Density Functions
- par(mfrow = c(2,2))
- # multiframe
- x<-0:12 #look at the first 12 probabilities
- plot (x, dpois(x, 2), xlab = "Number of Hits", ylab = "P(X = x)", type = "h", main= "Web Site Hits: Poisson(2)")
- plot (x, dpois(x, 3), xlab = "Number of Calls", ylab = "P(X = 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)")

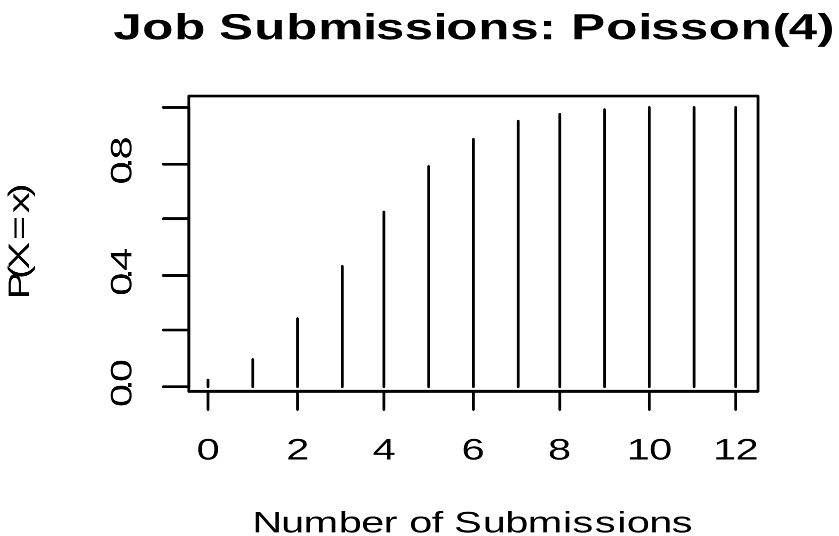


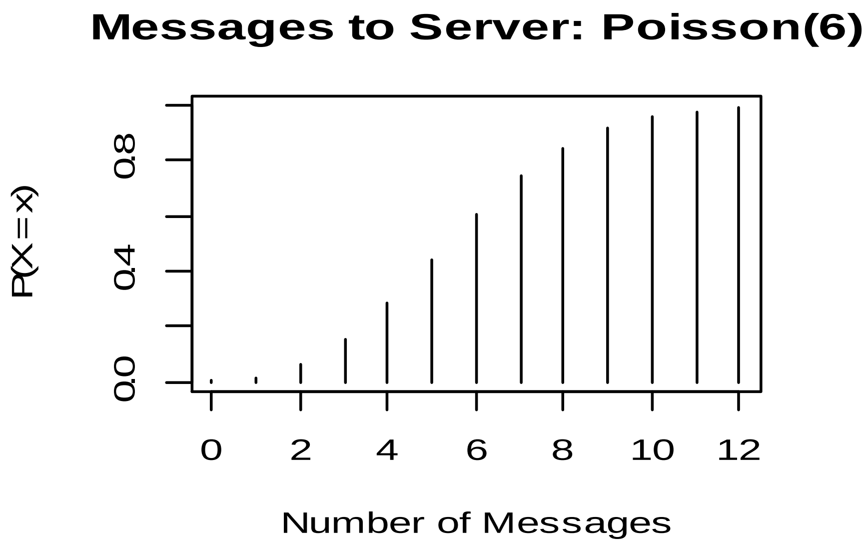
Web Site Hits: Poisson(2)
P(X = x)
0.00
0.10
0.20
0
2
4
6
8
10
12
Number of Hits
Calls to Mobile: Poisson(3)
P(X = x)
0.00
0.10
0.20
0
2
4
6
8
10
12
Number of Calls
Job Submissions: Poisson(4)
P(X = x)
0.00
0.10
0.20
0
2
4
6
8
10
12
Number of Submissions
Messages to Server: Poisson(6)
P(X = x)
0.00
0.10
0
2
4
6
8
10
12
Number of Messages

- Poisson Cumulative Distribution 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)")


Web Site Hits: Poisson(2)
P(X=x)
0.2
0.6
1.0
0 2 4 6 8 10 12
Number of Hits


Calls to Mobile: Poisson(3)
P(X=x)
0.2
0.6
1.0
0 2 4 6 8 10 12
Number of Calls


Job Submissions: Poisson(4)
P(X=x)
0.0
0.4
0.8
0 2 4 6 8 10 12
Number of Submissions


Messages to Server: Poisson(6)
P(X=x)
0.0
0.4
0.8
0 2 4 6 8 10 12
Number of Messages

› Practice Problems:-( Binomial and Poisson distribution)

1. For a random variable X with a binomial (20,1/2) distribution, find the following probabilities.

   (i). Find Pr(X < 8)

   (ii). Find Pr(X > 12)

   (iii). Find Pr(8 <= X <= 1

2) Let X be the number of heads in 10 tosses of a fair coin.

   1. Find the probability of getting at least 5 heads (that is, 5 or more).
   2. Find the probability of getting exactly 5 heads.
   3. Find the probability of getting between 4 and 6 heads, inclusive

3. A recent national study showed that approximately 55.8% of college students have used Google as a source in at least one of their term papers. Let X equal the number of students in a random sample of size n = 42 who have used Google as a source:

- 1. How is X distributed?
- 2. Sketch the probability mass function (roughly).
- 3. Sketch the cumulative distribution function (roughly).
- 4. Find the probability that X is equal to 17.
- 5. Find the probability that X is at most 13.
- 6. Find the probability that X is bigger than 11.
- 7. Find the probability that X is at least 15.
- 8. Find the probability that X is between 16 and 19, inclusive
- 9. Give the mean of X, denoted IE X.
- 10. Give the variance of X.
- 11. Give the standard deviation of X. 1
- 12. Find IE(4X + 51:324) 13. Compare mean and variance