ISI-BUDS
PROBABILITY - PART II

Veronica J. Berrocal

July 11, 2022

PART II PLAN:

- Random variables
- Some common distributions for discrete random variables

MOVING FROM EVENTS/OUTCOMES TO RANDOM VARIABLES

- We have seen the definition of random circumstances.
- In several instances, we assign a number to each outcomes of a random circumstance.
- *The mapping or function that assigns a number to each outcome of a random circumstance* is called a random variable (r.v.).

EXAMPLE: PEOPLE WITH A CERTAIN BLOOD TYPE

- Suppose we are selecting 10 students at random from this class. The number of students in the subset with blood type O is a random variable. The values that this random variable can take are: …. $\{0, 1, 2, \ldots 10\}$

DISCRETE VS CONTINUOUS RANDOM VARIABLES

- There are two different types of random variables:
  - Discrete random variables: that is, random variables that can take one of a countable list of distinct values
  - Continuous random variables: that is, random variables that can take any value in an interval or collection of intervals.

- Note: A discrete random variable can also take an infinite number of values. A countable list of values is such that the values can be counted one at a time, even if the counting never finishes.

DISCRETE OR CONTINUOUS?

A book is chosen at random from a library shelf. For each of the following characteristics of the book, decide whether the characteristic is a continuous or discrete random variable:

- Weight of the book (in lbs) continuous
- Number of chapters in the book discrete
- Width of the book continuous
- Book type (paperback, hardback) discrete
- Number of typos in the book discrete

WHY THE DIFFERENCE?

- Why do we differentiate between discrete and continuous random variables?
- For discrete random variables, we can calculate probabilities for exact outcomes.
- For continuous random variables, we cannot find probabilities for exact outcomes. We can only find probabilities for interval of values.

EXAMPLE: WAITING ON STANDBY

Suppose you are on the standby list to board a flight and you are the first on the list. You are interested in the following random variables:

1. the number of standby passengers that will be allowed to board the plane;
2. the flying time.

Suppose you want to know the following probabilities:

(a) The probability that at least one standby passenger is allowed to board the plane;

(b) The probability that the flying time is equal to the time specified in the flight schedule online. X not possible

Can you calculate both probabilities?

NOTATIONS FOR DISCRETE RANDOM VARIABLES

- We will use the following notation when specifying the probabilities associated with a discrete random variable.
- We denote with:
  - X the random variable;
  - k a specified value that the the random variable X could assume;
  - $P(X=k)$: the probability that the random variable X is equal to k.

PROBABILITY DISTRIBUTION FOR A DISCRETE R.V.

$P(X=k)$ for any value k that X takes

- The probability distribution function (pdf) for a discrete random variable X is a table or a rule that assigns probabilities to the possible values of the random variable X.

- Note: sometimes the probability distribution function for discrete random variables is also called probability mass function (pmf).

k	0	1	2	...
$P(X=k)$				

$P(X=k)$ = function that :values k

EXAMPLE: NUMBER OF COURSES TAKEN BY STUDENTS

Suppose that 35% of students at "Happy University" typically take 4 courses during a semester, 45% take 5 courses and the remaining 20% take 6 courses. Let X be the number of courses a randomly selected student at "Happy University" takes.

What is the probability distribution function for X?

k	0	1	2	3	4	5	6
$P(X=k)$	0	0	0	0	.35	.45	.2

EXAMPLE: NUMBER OF GIRLS

$X=1 \rightarrow$ GBB, BGB, BBG

$X=2$ BGG, GBG, GGB

Suppose that a family has 3 children, and that the probability of a girl is ½ for each birth.

Let X be the random variable "number of girls among the three children". Then, the possible values of X are 0, 1, 2 and 3.

What is the probability distribution function for X?

k	0	1	2	3
$P(X=k)$	$(1/2)^3$	$3\cdot(1/2)^3$	$3\cdot(1/2)^3$	$(1/2)^3$

CUMULATIVE DISTRIBUTION FUNCTION

- A cumulative distribution function (cdf) for a discrete random variable X is a table or a rule that provides $P(X \leq k)$ for any value of k, that is, the probability that the random variable X is less or equal to a specific value k.

- Example: Consider the example of a family with 3 children presented earlier.

TRUE OR FALSE: The probability that two or fewer of the children are girls can be obtained from the cumulative distribution function.

$F(2) = P(X \leq 2)$ Value of the cdf at 2

FINDING PROBABILITIES FOR COMPLEX EVENTS

- Knowing the probability distribution function of a random variable X allows us to find probabilities for complex events easily.
- Consider again the example of the family with 3 children. What is the probability that the family will have at least one child of each sex? Remember X is the random variable "number of girls in 3 children".

$$P(X = 1 \text{ or } X = 2) = P(X=1) + P(X=2)$$

EXPECTED VALUE OF A RANDOM VARIABLE

- The expected value of a random variable X is the mean value that the random variable will obtain in an infinite number of observations of the random variable.
- It is denoted by E(X) and it is calculated as:

$$\boxed{\mathrm{E(X)} = \sum_{all\ possible\ values\ of\ k} k * P(X = k)}$$

EXAMPLE: GAMBLING LOSSES

Suppose that in a gambling game, the probability of winning \$2 is 0.3, and the probability of losing \$1 is 0.7. Let X denote the amount a player "gains" on a single play.

What is the expected amount a player will "gain" per bet over many plays?

x	2	-1
$P(X=x)$	0.3	0.7

$$E(X) = (2 * 0.3) + (-1 * 0.7) = -0.1$$

VARIANCE OF A DISCRETE RANDOM VARIABLE

- The variance (and the standard deviation) of a discrete random variable X quantifies how spread the possible values of a discrete random variable might be, weighted by how likely each value is to occur.
- The variance of a discrete random variable X is denoted with Var(X) and can be derived via:

$$\boxed{\mathrm{Var(X)} = \sum_{all\ possible\ values\ of\ k} (k-\mu)^2 * P(X=k)}$$

Where $\mu = E(X)$.

EXAMPLE:

Suppose you decide to invest $100 in a scheme that you hope will make some money. You have two choices of investment plans, and you must decide which one to choose. For each $100 invested under the two plans, the possible net gains one year later and their probabilities are shown.

Plan 1		Plan 2	
Net Gain	Probability	Net Gain	Probability
$5,000	0.001	$20	0.3
$1000	0.005	$10	0.2
$0	0.994	$4	0.5

Which one will you choose?

BINOMIAL RANDOM VARIABLES

- An important class of discrete random variables are the Binomial random variables.
- In order for a discrete random variable X to be a Binomial random variable, the following conditions must hold:
  - X must provide the number of times a given event occurs in repeated trials
  - The trials must be: independent, and allow only two possible outcomes (what we call a "*success*", and what we call a "*failure*").
  - The probability of "*success*" in any given trial should remain always the same.

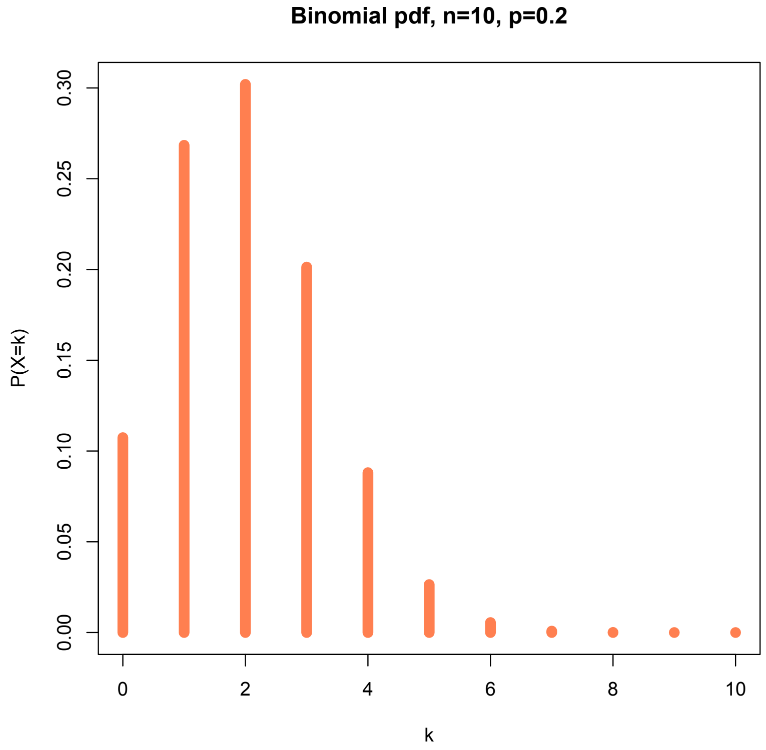
BINOMIAL P.D.F.

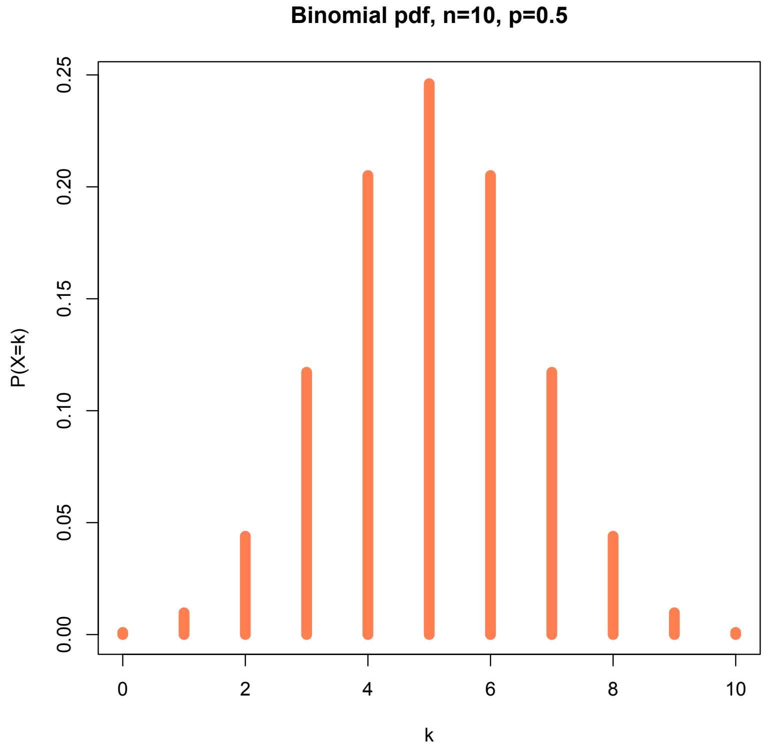
- For a Binomial random variable X, the probabilities for the possible value of X are given by the following probability distribution function:

$$\boxed{P(X=k)=\frac{n!}{k!\,(n-k)!}p^k\cdot(1-p)^{n-k}}$$

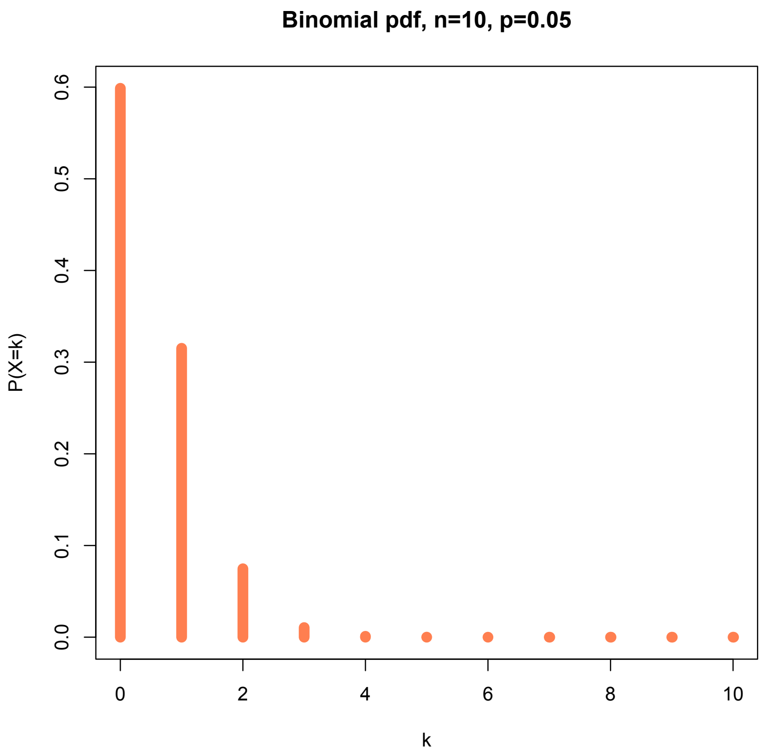
For k=0, 1, 2,..., n.

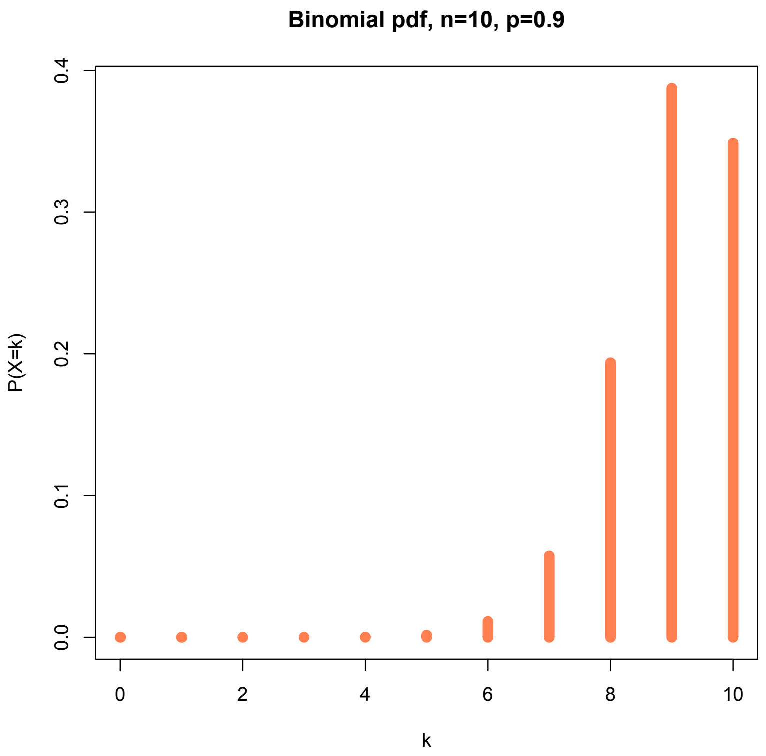
BINOMIAL P.D.F.





BINOMIAL P.D.F.





BINOMIAL R.V.'S: EXPECTED VALUE AND VARIANCE

- If X is a Binomial random variable, then

$$\boxed{E(X) = n \times p}$$

$$\boxed{Var(X) = n * p * (1-p)}$$

where $p=P(X=1)$ or probability of success and n is the number of trails.

EXAMPLE: PASSING A QUIZ BY GUESSING

You have been busy lately, so busy that you are surprised by a 15-questions TRUE-FALSE quiz given in today's statistics class. The quiz is about reading that you have not done, so you are forced to guess at every question. You will pass the quiz if you get 10 or more correct answers.

If we call X the random variable indicating the number of correct answers, how do you calculate the probability that you pass the quiz?

$X \sim \text{Binomial}(n=15, p=0.5)$

$P(\text{passing}) = P(X \geq 10) = 1 - P(X \leq 9)$

$= 1 - \text{pbinom}(9, 15, 0.5)$

NEGATIVE BINOMIAL RANDOM VARIABLES

- Another important class of discrete random variables are the Negative Binomial random variables.
- In order for a discrete random variable X to be a Negative Binomial random variable, the following conditions must hold:
  - We are performing independent trials, each allowing only two possible outcomes (what we call a "*success*", and what we call a "*failure*"), and each trial has the same probability p of success;
  - X counts the number of failures we see before we obtain r successes.

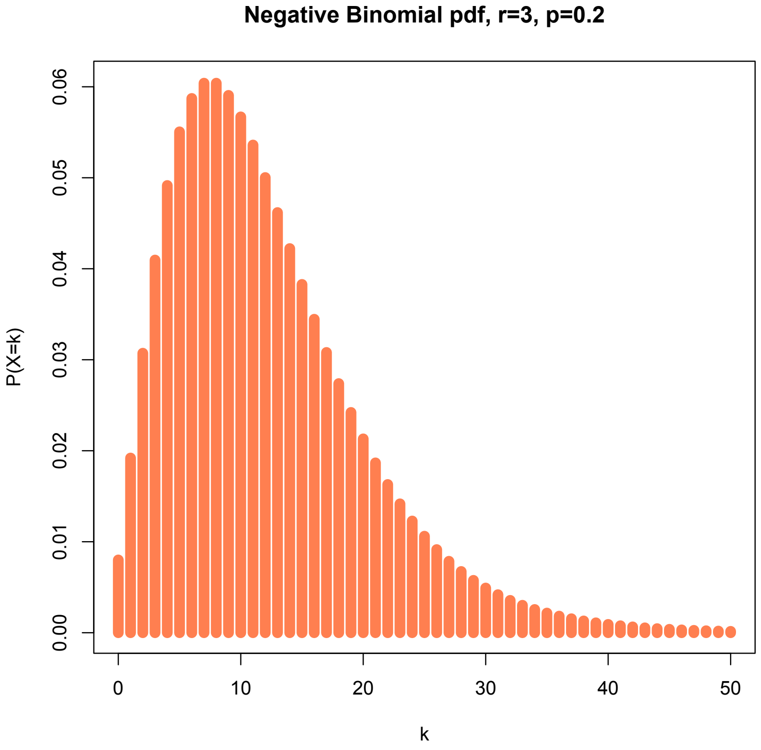
NEGATIVE BINOMIAL P.D.F.

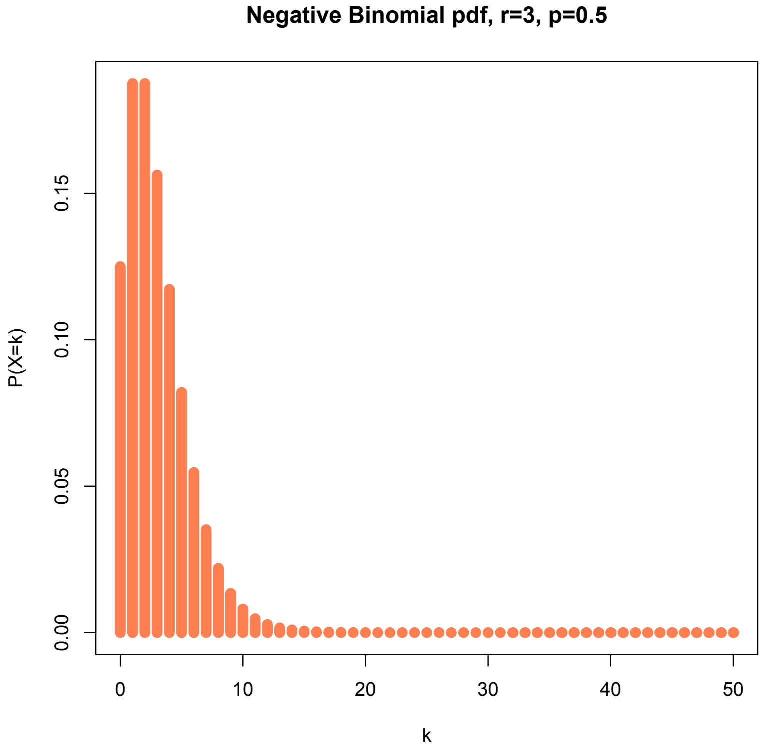
- For a Negative Binomial random variable X, the probabilities for the possible value of X are given by the following probability distribution function:

$$\boxed{P(X=k)=\frac{(k+r-1)!}{(r-1)!\,k!}p^r\cdot(1-p)^k}$$

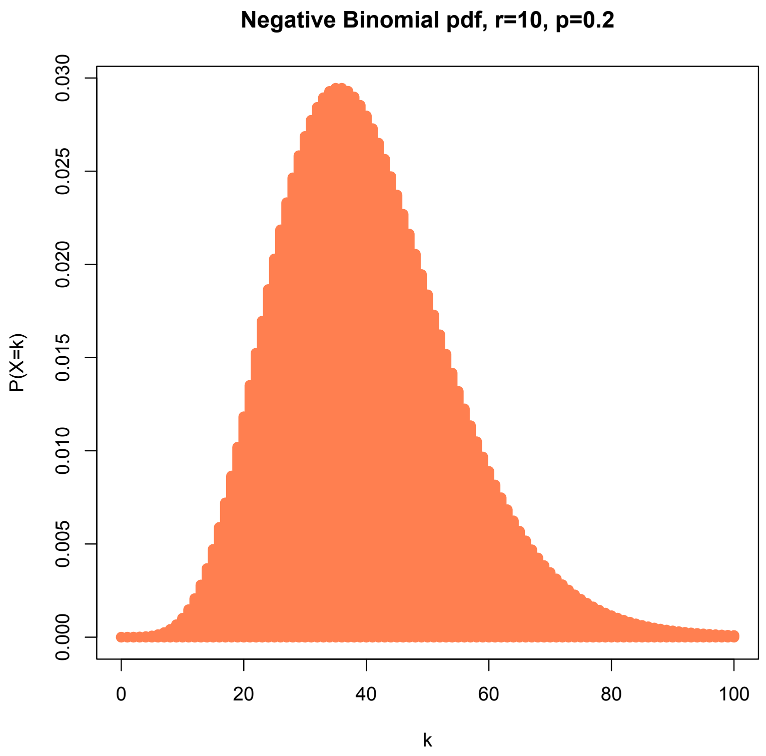
For $k=0, 1, 2,\ldots,\ldots$

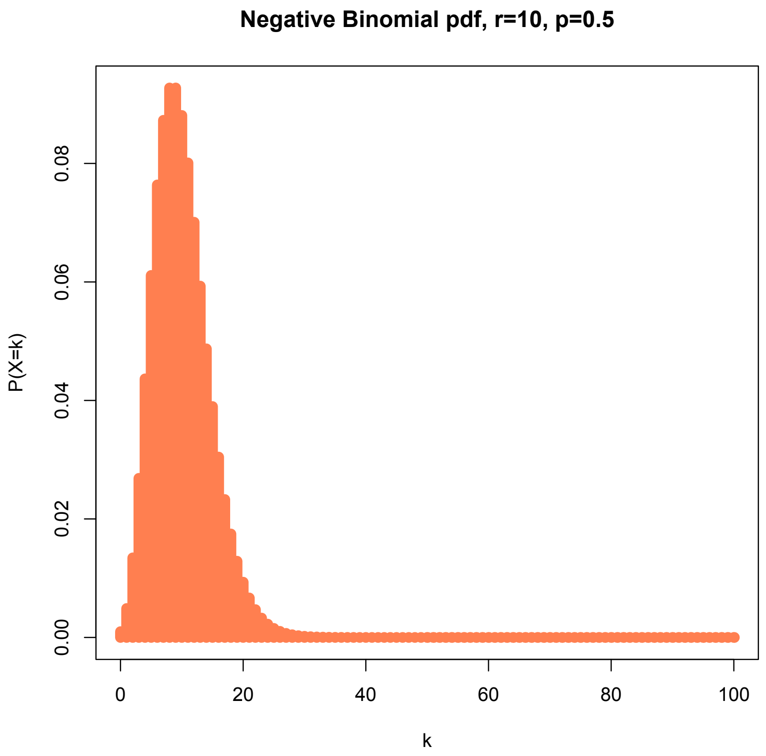
NEGATIVE BINOMIAL P.D.F.





NEGATIVE BINOMIAL P.D.F.





NEGATIVE BINOMIAL R.V.'S: EXPECTED VALUE AND VARIANCE

- If X is a Negative Binomial random variable, then

$$\boxed{E(X) = \frac{(1-p)*r}{p}}$$

$$\boxed{Var(X) = \frac{(1-p)*r}{p^2}}$$

where p=P(X=1) or probability of success and r is the number of successes that we want to achieve.

EXAMPLE: WINNING THE LOTTERY?

The probability of winning in a certain state lottery is stated to be 1/9.

What is the expected number of non-winning tickets that one will have purchased in order to get to two wins in this state lottery?

POISSON RANDOM VARIABLES

- The last class of discrete random variables that we will consider are called Poisson random variables.
- In order for a discrete random variable X to be a Poisson random variable, the following condition must hold:
  - X counts the number of events occurring in a given interval of time (or in a subset of space).
  - Note: There is a relationship between Poisson and Binomial random variables.

POISSON P.D.F.

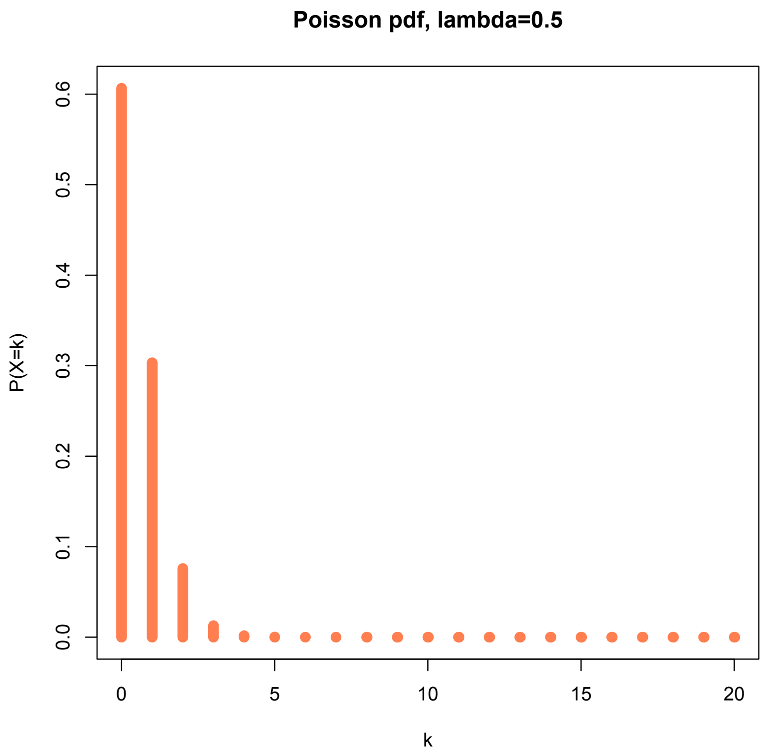
- For a Poisson random variable X, the probabilities for the possible value of X are given by the following probability distribution function:

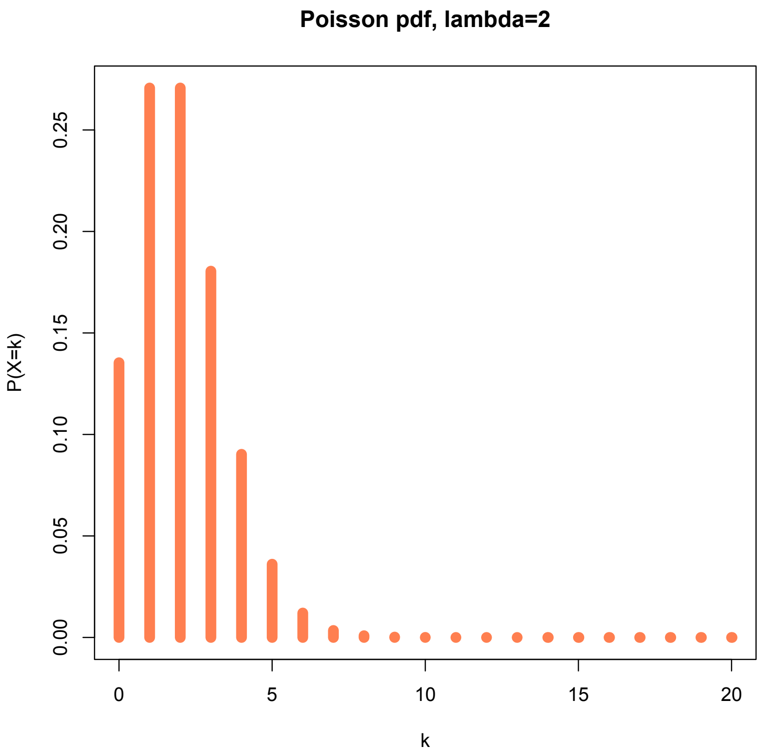
$$\boxed{P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}}$$

$\lambda > 0$

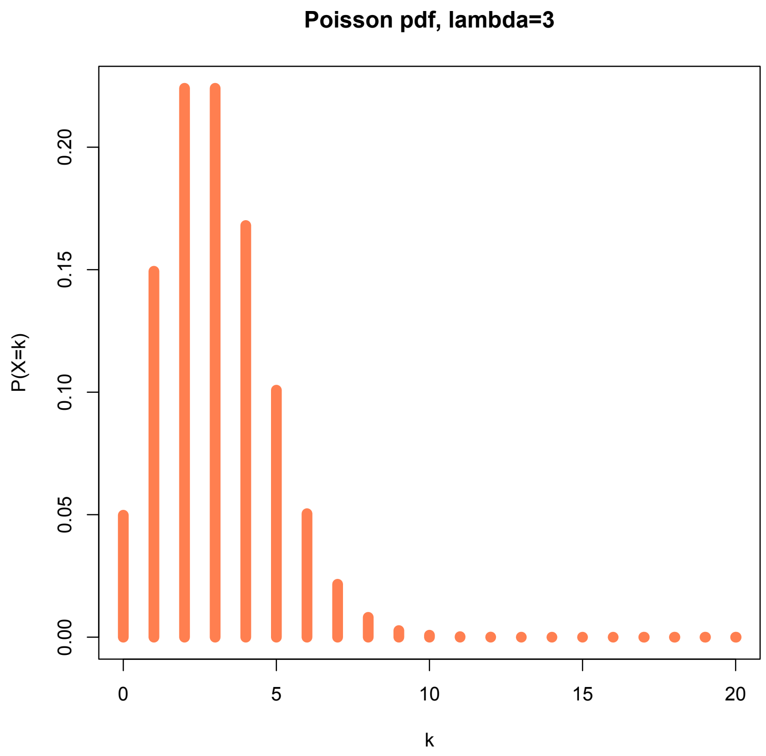
where λ is called the parameter of the distribution, and k=0, 1, 2,...,...

POISSON P.D.F.

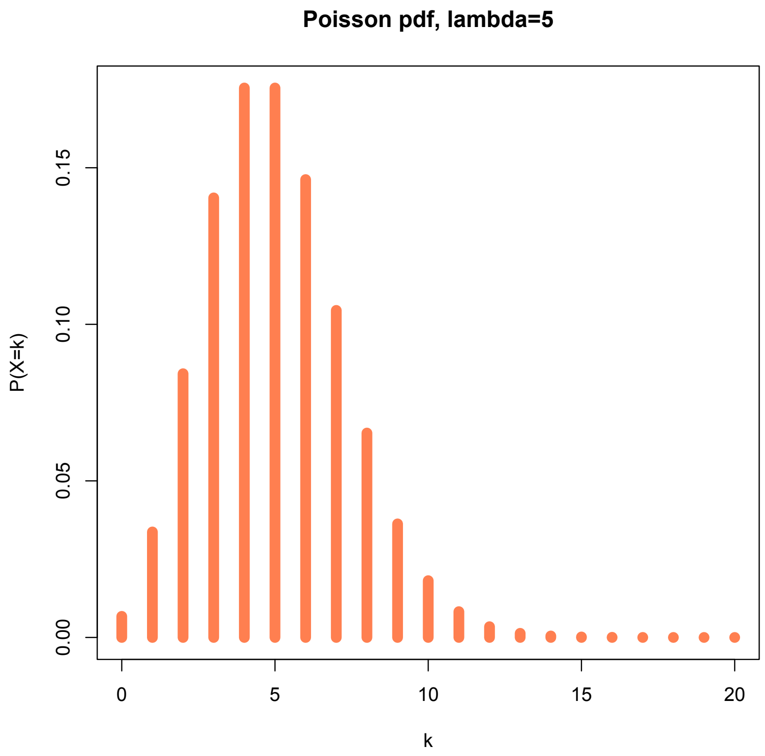

Poisson pdf, lambda=0.5
P(X=k)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
5
10
15
20
k


Poisson pdf, lambda=2
P(X=k)
0.00
0.05
0.10
0.15
0.20
0.25
0
5
10
15
20
k

POISSON P.D.F.



Poisson pdf, lambda=3
P(X=k)
0.00
0.05
0.10
0.15
0.20
0
5
10
15
20
k



Poisson pdf, lambda=5
P(X=k)
0.00
0.05
0.10
0.15
0
5
10
15
20
k

POISSON R.V.'S: EXPECTED VALUE AND VARIANCE

- If X is a Poisson random variable with parameter λ, then

$$\boxed{E(X) = \lambda} \qquad \boxed{Var(X) = \lambda.}$$

When $Var(X) > E(X)$ you have an overdispersed Poisson distribution/r.v.

EXAMPLE: OFFICE CALLS

Suppose that an office receives telephone calls at a rate of 0.5 calls per minute.

If we consider the number X of calls in a 5-minutes interval, then X follows a Poisson distribution with parameter $\lambda = 2.5.$

$$P(X=0) = \frac{e^{-2.5} \cdot (2.5)^{0}}{0!} = \frac{e^{-2.5} \cdot 1}{0!}$$

$$0! = 1$$

What is the probability that the office receives no call in a 5-minutes interval? What is the probability that the office will receive exactly one call?