CS 107, Probability, Spring 2019
Lecture 23

Michael Poghosyan

AUA

18 March 2019

Content

- Examples of Important Discrete R.V.s

LZ

Let us prove by Induction that **all horses are the same color**

LZ

Let us prove by Induction that **all horses are the same color**

Figure: Same color horses ¨⌣

LZ

Statement: All horses are the same color.

Statement: All horses are the same color.
Proof: By induction. First we write the statement as:

LZ

Statement: All horses are the same color.
Proof: By induction. First we write the statement as:
For any natural number n, in any set of n horses, all horses in that set have the same color.

LZ

Statement: All horses are the same color.
Proof: By induction. First we write the statement as:
For any natural number n, in any set of n horses, all horses in that set have the same color.

- Clearly, the statement is true for $n = 1$.

Statement: All horses are the same color.
Proof: By induction. First we write the statement as:
For any natural number n, in any set of n horses, all horses in that set have the same color.

- Clearly, the statement is true for $n = 1$.
- Assume now that the statement is true for n. Let us prove for $n + 1$. We consider a set of $(n + 1)$ horses: $\{h_1, h_2, ..., h_{n+1}\}$.

Statement: All horses are the same color.
Proof: By induction. First we write the statement as:
For any natural number n, in any set of n horses, all horses in that set have the same color.

- Clearly, the statement is true for $n = 1$.
- Assume now that the statement is true for n. Let us prove for $n + 1$. We consider a set of $(n + 1)$ horses: $\{h_1, h_2, ..., h_{n+1}\}$. We split this set into two size-n subsets:

$$\underbrace{h_1, \overbrace{h_2, h_3, ..., h_{n-1}, h_n}^{\text{n horses}}}_{\text{n horses}}, h_{n+1}$$

LZ

Statement: All horses are the same color.
Proof: By induction. First we write the statement as:
For any natural number n, in any set of n horses, all horses in that set have the same color.

- Clearly, the statement is true for $n = 1$.
- Assume now that the statement is true for n. Let us prove for $n + 1$. We consider a set of $(n + 1)$ horses: $\{h_1, h_2, \dots, h_{n+1}\}$. We split this set into two size-n subsets:

$$\underbrace{h_1, \rlap{$\overbrace{\phantom{h_2, h_3, \dots, h_{n-1}, h_n, h_{n+1}}}^{\text{n horses}}$}h_2, h_3, \dots, h_{n-1}, h_n}_{\text{n horses}}, h_{n+1}$$

By the induction assumption, horses in each set have the same color. Hence, all $(n + 1)$ horses will have the same color!

LZ

Discrete Uniform Distribution

Recall the definition of the Discrete Uniform Distribution:

Discrete Uniform Distribution

We say that the r.v. X has a Discrete Uniform Distribution with (over) the values $x_1, x_2, \ldots, x_n$ ($x_i \neq x_j$, $i \neq j$), and we will write $X \sim DiscreteUnif(x_1, \ldots, x_n)$, if

Values of X	x_1	x_2	$\ldots$	x_n
$\mathbb{P}(X = x)$	$\frac{1}{n}$	$\frac{1}{n}$	$\ldots$	$\frac{1}{n}$

Discrete Uniform Distribution

Recall the definition of the Discrete Uniform Distribution:

Discrete Uniform Distribution

We say that the r.v. X has a Discrete Uniform Distribution with (over) the values $x_1, x_2, ..., x_n$ $(x_i \neq x_j,\ i \neq j)$, and we will write $X \sim DiscreteUnif(x_1, ..., x_n)$, if

Values of X	x_1	x_2	...	x_n
$\mathbb{P}(X = x)$	$\frac{1}{n}$	$\frac{1}{n}$	...	$\frac{1}{n}$

that is, $\mathbb{P}(X = x_k) = \dfrac{1}{n}$, for any $k = 1, 2, ..., n$.

Discrete Uniform Distribution: R Examples

R Code

```
#Discrete Uniform on data
data <- c(-1,2,4)
sample(data, size = 10, replace = T)

#Generating a sample of size 100
s <- sample(data, size = 100, replace = T)
#Calculating the number of -1 in the sample s
length(s[s == -1])
```

Bernoulli Distribution

Bernoulli Distribution

We will say that the r.v. X has a Bernoulli Distribution with the parameter (probability) $p \in [0,1]$, and we will write $X \sim \mathit{Bernoulli}(p)$, if it has the following PMF:

Values of X	0	1
$\mathbb{P}(X = x)$	$1 - p$	p

Bernoulli Distribution

Bernoulli Distribution

We will say that the r.v. X has a Bernoulli Distribution with the parameter (probability) $p \in [0, 1]$, and we will write $X \sim Bernoulli(p)$, if it has the following PMF:

Values of X	0	1
$\mathbb{P}(X = x)$	$1 - p$	p

that is, $\mathbb{P}(X = 1) = p$, and $\mathbb{P}(X = 0) = 1 - p$.

Bernoulli Distribution

Bernoulli Distribution

We will say that the r.v. X has a Bernoulli Distribution with the parameter (probability) $p \in [0,1]$, and we will write $X \sim \mathit{Bernoulli}(p)$, if it has the following PMF:

Values of X	0	1
$\mathbb{P}(X = x)$	$1-p$	p

that is, $\mathbb{P}(X=1) = p$, and $\mathbb{P}(X=0) = 1-p$.
The same can be written in the form (for[a] $p \in (0,1)$):

$$\mathbb{P}(X = x) = p^x \cdot (1-p)^{1-x}, \quad x \in \{0,1\}.$$

[a] To avoid numbers like 0^0.

Bernoulli Distribution: Examples

Bernoulli Distribution models the binary outcomes case - Success/Failure, Yes/No, Boy/Girl, Pass/Fail, Healthy/Ill, Smoker/Non-Smoker, Flan/Fstan, ... Some examples:

Bernoulli Distribution: Examples

Bernoulli Distribution models the binary outcomes case - Success/Failure, Yes/No, Boy/Girl, Pass/Fail, Healthy/Ill, Smoker/Non-Smoker, Flan/Fstan, ... Some examples:

- We are tossing a fair coin, and $X = 0$ if Heads appears, and $X = 1$ in the other case. Then $X \sim Bernoulli\left(\frac{1}{2}\right)$;

Bernoulli Distribution: Examples

Bernoulli Distribution models the binary outcomes case - Success/Failure, Yes/No, Boy/Girl, Pass/Fail, Healthy/Ill, Smoker/Non-Smoker, Flan/Fstan, ... Some examples:

- We are tossing a fair coin, and $X = 0$ if Heads appears, and $X = 1$ in the other case. Then $X \sim \textit{Bernoulli}\left(\frac{1}{2}\right)$;
- We are rolling a fair die. $X = 0$, if the shown number is > 4, and $X = 1$ otherwise. Then $X \sim \textit{Bernoulli}(\quad)$;

Bernoulli Distribution: Examples

Bernoulli Distribution models the binary outcomes case - Success/Failure, Yes/No, Boy/Girl, Pass/Fail, Healthy/Ill, Smoker/Non-Smoker, Flan/Fstan, ... Some examples:

- We are tossing a fair coin, and $X = 0$ if Heads appears, and $X = 1$ in the other case. Then $X \sim \mathit{Bernoulli}\left(\frac{1}{2}\right)$;
- We are rolling a fair die. $X = 0$, if the shown number is > 4, and $X = 1$ otherwise. Then $X \sim \mathit{Bernoulli}(\quad)$;
- Say, $X = 1$ in case the (randomly chosen Prob) student will pass the course, and $X = 0$ otherwise. Say, we can model X as a Bernoulli r.v.: $X \sim \mathit{Bernoulli}(0.76)$;

Bernoulli Distribution: Examples

Bernoulli Distribution models the binary outcomes case - Success/Failure, Yes/No, Boy/Girl, Pass/Fail, Healthy/Ill, Smoker/Non-Smoker, Flan/Fstan, ... Some examples:

- We are tossing a fair coin, and $X = 0$ if Heads appears, and $X = 1$ in the other case. Then $X \sim Bernoulli\left(\frac{1}{2}\right)$;
- We are rolling a fair die. $X = 0$, if the shown number is > 4, and $X = 1$ otherwise. Then $X \sim Bernoulli(\quad)$;
- Say, $X = 1$ in case the (randomly chosen Prob) student will pass the course, and $X = 0$ otherwise. Say, we can model X as a Bernoulli r.v.: $X \sim Bernoulli(0.76)$;
- We are interested if a specific student will get the $A+$ grade from Probability: we can take $X = 1$ to indicate that he/she will get $A+$, and $X = 0$ otherwise. Then $X \sim Bernoulli(p)$, and p depends on the student: if he/she is working well, then p is close to 1.

Bernoulli Distribution: Examples

- We are interested if the patient has a specific disease or not: we can take, for example, $X = 1$, if he/she has that disease, and $X = 0$ otherwise. p will be the probability of having that disease, and $X \sim Bernoulli(p)$;

Bernoulli Distribution: Examples

- We are interested if the patient has a specific disease or not: we can take, for example, $X = 1$, if he/she has that disease, and $X = 0$ otherwise. p will be the probability of having that disease, and $X \sim Bernoulli(p)$;
- We are interested if the nearest toilet room is busy at this moment $\ddot\smile$: if it is, we can denote that by $X = 1$, and $X = 0$ otherwise. Then $X \sim Bernoulli(p)$, but I do not have any ideas about p;

Bernoulli Distribution: Examples

- We are interested if the patient has a specific disease or not: we can take, for example, $X = 1$, if he/she has that disease, and $X = 0$ otherwise. p will be the probability of having that disease, and $X \sim Bernoulli(p)$;
- We are interested if the nearest toilet room is busy at this moment $\ddot\smile$: if it is, we can denote that by $X = 1$, and $X = 0$ otherwise. Then $X \sim Bernoulli(p)$, but I do not have any ideas about p;
- An insurance company is interested if the claim size for today will exceed 10Mln AMD: we can take $X = 1$, if it will exceed and $X = 0$ if not.

Bernoulli Distribution: Examples

- We are interested if the patient has a specific disease or not: we can take, for example, $X = 1$, if he/she has that disease, and $X = 0$ otherwise. p will be the probability of having that disease, and $X \sim Bernoulli(p)$;
- We are interested if the nearest toilet room is busy at this moment $\ddot\smile$: if it is, we can denote that by $X = 1$, and $X = 0$ otherwise. Then $X \sim Bernoulli(p)$, but I do not have any ideas about p;
- An insurance company is interested if the claim size for today will exceed 10Mln AMD: we can take $X = 1$, if it will exceed and $X = 0$ if not.
- We are interested if the max stock price for FB next month will be higher than the stock price today: we can take $X = 1$, if that max price next month will be larger, and $X = 0$ otherwise.

Bernoulli Distribution: Examples

- We are interested if some specific Football team will win the next game: $X = 1$ means will win, and $X = 0$ otherwise. Then $X \sim \mathit{Bernoulli}(p)$;
- We are interested if some specific Football team will win the next game with the difference in the scores more than 3: $X = 1$ if that will be the case, and $X = 0$ otherwise. Then $X \sim \mathit{Bernoulli}(p)$;

Bernoulli Distribution: Examples

- We are interested if some specific Football team will win the next game: $X = 1$ means will win, and $X = 0$ otherwise. Then $X \sim Bernoulli(p)$;
- We are interested if some specific Football team will win the next game with the difference in the scores more than 3: $X = 1$ if that will be the case, and $X = 0$ otherwise. Then $X \sim Bernoulli(p)$;
- And you can make infinitely many examples of Bernoulli distributed r.v.s

Bernoulli Distribution: R Examples

R Code

```
#Bernoulli Distribution
p <- 0.3
#In fact, we use the Binomial Distr with size = 1
x <- rbinom(10, size = 1, prob = p)
x
```

Binomial Distribution

Binomial Distribution

We will say that the r.v. X has a Binomial Distribution with the parameters n and p, $p \in [0, 1]$, and we will write $X \sim Binom(n, p)$, if it has the following PMF:

Values of X	0	1	2	...	n
$\mathbb{P}(X = x)$	$\binom{n}{0} p^0 (1-p)^{n-0}$	$\binom{n}{1} p^1 (1-p)^{n-1}$	$\binom{n}{2} p^2 (1-p)^{n-2}$	...	$\binom{n}{n} p^n (1-p)^{n-n}$

Binomial Distribution

Binomial Distribution

We will say that the r.v. X has a Binomial Distribution with the parameters n and p, $p \in [0,1]$, and we will write $X \sim Binom(n,p)$, if it has the following PMF:

Values of X	0	1	2	...	n
$\mathbb{P}(X=x)$	$\binom{n}{0}p^0(1-p)^{n-0}$	$\binom{n}{1}p^1(1-p)^{n-1}$	$\binom{n}{2}p^2(1-p)^{n-2}$	...	$\binom{n}{n}p^n(1-p)^{n-n}$

that is (in the case $p \in (0,1)$),

$$\mathbb{P}(X=x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}, \qquad x = 0, 1, 2, ..., n$$

Binomial Distribution

Binomial Distribution

We will say that the r.v. X has a Binomial Distribution with the parameters n and p, $p \in [0,1]$, and we will write $X \sim Binom(n,p)$, if it has the following PMF:

Values of X	0	1	2	...	n
$\mathbb{P}(X=x)$	$\binom{n}{0}p^0(1-p)^{n-0}$	$\binom{n}{1}p^1(1-p)^{n-1}$	$\binom{n}{2}p^2(1-p)^{n-2}$	...	$\binom{n}{n}p^n(1-p)^{n-n}$

that is (in the case $p \in (0,1)$),

$$\mathbb{P}(X=x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}, \qquad x = 0,1,2,...,n$$

Exercise: Check that the sum of probabilities is 1.

Binomial Distribution: Interpretation

Binomial Distribution gives the model for the repeated trials.

Binomial Distribution: Interpretation

Binomial Distribution gives the model for the repeated trials.
If $X \sim Binom(n, p)$, then

Binomial Distribution: Interpretation

Binomial Distribution gives the model for the repeated trials.
If $X \sim Binom(n, p)$, then

- we are doing n independent repetitions of some Simple Experiment,

Binomial Distribution: Interpretation

Binomial Distribution gives the model for the repeated trials.
If $X \sim Binom(n, p)$, then

- we are doing n independent repetitions of some Simple Experiment,
- in each Simple Experiment some Event can happen (Success) with the Probability p;

Binomial Distribution: Interpretation

Binomial Distribution gives the model for the repeated trials.
If $X \sim Binom(n, p)$, then

- we are doing n independent repetitions of some Simple Experiment,
- in each Simple Experiment some Event can happen (Success) with the Probability p;
- X shows the number of Successes we will have during that n trials

Binomial Distribution: Examples

- We can toss a fair coin 12 times, and we are interested in how many heads we will get during that tosses. If we will denote by X the number of heads we will have, then $X \sim Binom(12, 0.5)$.

Binomial Distribution: Examples

- We can toss a fair coin 12 times, and we are interested in how many heads we will get during that tosses. If we will denote by X the number of heads we will have, then $X \sim Binom(12, 0.5)$.
- Assume we have 5 red and 10 green balls in the box. We are drawing a ball at random, fixing its color, returning it to the box, and then doing that again 100 times. We are interested in the number of green balls we will pick. If we will denote by X the number of green balls taken out of the box during this 100 trials, then $X \sim Binom(100, 2/3)$, since the probability to get a green ball in each trial is $p = \frac{10}{15} = \frac{2}{3}$ (we assume drawing each ball has the same probability for all balls).
- Your examples?

Binomial Distribution: R Examples

R Code

```
#Binomial Distribution
p <- 0.45
#Size is our n
x <- rbinom(20, size = 3, prob = p)
x

#Plotting the PMF
size = 21
x <- 0:size
PMF <- dbinom(x, size = size, prob = p)
plot(x,PMF, pch = 19)
```

Supplementary: Very short R Intro

Supplementary: Very short R Intro

Note: Please run the commands below to see what they are doing $\ddot\smile$

Basic Commands of **R**

Some assignments (The Sharp(Diez $\ddot{\smile}$, $\#$) sign denotes that after it the whole line is a comment):

```
x <- 10 # x=10
x = 10  # the same as above, x=10
x
print(x)
y <- 20
z <- x+y
z
3 -> x
```

Vector Arithmetics in **R**

Vector Arithmetics in **R**

```
x = c(2,3,5,7,2,1) # c means concatenation,
        #so x is an array with elements 2,3,...
y = c(10,15,12)
z = c(x,y) concatenates, joins the arrays
x
x[2]  #the second element in x
y
z
z^2 #squares the elements of z
x*x #squares the elements of x
```

Try now:

```
x+y
```

Can you guess what is doing **R** in this case?

Vector Arithmetics in R

Ways to create some regular-pattern vectors

```
x <- 1:5  #The same as x <- c(1,2,3,4,5)
x <- seq(from = 1, to = 10, by = 2)
x <- seq(from = 1, to = 10, length = 3)
y <- rep(x,2) #repeats (replicates) x 2 times
y <- rep(x, each = 2)
```

Vector Arithmetics in R

Ways to create some regular-pattern vectors

```
x <- 1:5  #The same as x <- c(1,2,3,4,5)
x <- seq(from = 1, to = 10, by = 2)
x <- seq(from = 1, to = 10, length = 3)
y <- rep(x,2) #repeats (replicates) x 2 times
y <- rep(x, each = 2)
```

Operations with vectors:

```
c(1,2,3)+c(2,1,3)
c(1,2,3)*c(2,1,3) #term-by-term multiplication
c(1,2,3)+1 #Adds 1 to every coordinate
1/c(1,2,3) #1/(every element)
c(1,2,3,4)+c(10,100) # repeats 10, 100 twice and
                     # adds to {1,2,3,4}
```

Vector Arithmetics in R

Choosing elements of a vector:

Vector Arithmetics in R

Choosing elements of a vector:

```
x <- c(2, -1, 9, 0 , 4, 1)
x[3] #the third element of x
x[2:4] #elements no. 2,3,4, of x
x[c(2,4,5)] #elements no. 2,4,5 of x
x > 0
x[x > 0] #Selects all positive elements of x
x[-c(2,3)] #x, without elements no 2,3
x[-{2:4}]  #x, without elements no 2,3,4
x[-(2:4)]  #x, without elements no 2,3,4
x[seq(from = 2, to = length(x), by = 2)]
        # will give the even-indexed elements of x
```

Vector Arithmetics in R

```
x <- c(2, -1, 9, 0 , -4, 1, 2)
x[x < 0] <- 0 #assigns 0 to all negative elements of x
which(x == 2) #gives the index of elements equal to 2
which(x > 0) #gives the indices of positive elements
which(x == max(x)) #returns the index (indices) of
                   #max element in x
which.max(x) #same thing as above
which.min(x)
```

Defining Functions

```
square = function(x) x^2
square(3)
square
```

Defining Functions

```
square = function(x) x^2
square(3)
square
```

The usage is:

```
function(arguments) body
```

Defining Functions

```
square = function(x) x^2
square(3)
square
```

The usage is:

```
function(arguments) body
```

```
f <- function(x) {x+1}
f(4)
f <- function(x) {return(x+1)}
f(4)
f <- function(x,y) {x*y}
f(2,3)
f <- function(x,y=10) {x*y}
f(2,3); f(2)
```

Plotting 2D graphs

```
square = function(x) x^2
plot(square)
plot(square, -1,3)
plot(square, xlim = c(-2,3))
plot(square, lwd=3, xlim = c(-2,3))
plot(square, lwd=3, col="red", xlim = c(-2,3))
plot(square, lwd=3, xlim = c(-2,3), ylim = c(0,1))
```

Plotting 2D graphs

```
square = function(x) x^2
plot(square)
plot(square, -1,3)
plot(square, xlim = c(-2,3))
plot(square, lwd=3, xlim = c(-2,3))
plot(square, lwd=3, col="red", xlim = c(-2,3))
plot(square, lwd=3, xlim = c(-2,3), ylim = c(0,1))
```

Another method is

```
curve(sin, -pi, pi)
curve(sin(x), -pi, 2*pi)
curve(sin, -pi, 2*pi, lwd = 4)
f <- function(x) {x+2*x^2-4*x^3}
curve(f, -2,2)
```

Vector Functions in R

```
x <- c(2,2,3,4,3,5,7,2,1)
max(x)
min(x)
range(x)
length(x)
sum(x)
prod(x)
cumsum(x)
cumprod(x)
```

Probability Distributions in **R**

The following names are used in **R** for distributions:

Probability Distributions in **R**

The following names are used in **R** for distributions:

Distribution	R name		Distribution	R name
Beta	beta		Lognormal	lnorm
Binomial	binom		Negative Binomial	nbinom
Cauchy	cauchy		Normal	norm
Chisquare	chisq		Poisson	pois
Exponential	exp		Student t	t
F	f		Uniform	unif
Gamma	gamma		Tukey	tukey
Geometric	geom		Weibull	weib
Hypergeometric	hyper		Wilcoxon	wilcox
Logistic	logis			

Probability Distributions in R

Usage:

Probability Distributions in **R**

Usage:

Name	Description
dname( )	PD(M)F of distribution
pname( )	CDF
qname( )	quantile function
rname( )	random numbers generated from distribution

Probability Distributions in R

Usage:

Name	Description
dname( )	PD(M)F of distribution
pname( )	CDF
qname( )	quantile function
rname( )	random numbers generated from distribution

Example:

```
rnorm(10)
dunif(1)
pexp(3, rate = 2)-pexp(-3, rate = 2)
qchisq(0.3, df = 1)
```

Probability Distributions in R

Note that the distributions above have some parameters (say, min and max for Uniform, mean and sd for Normal etc.), and some parameters have default values you can keep (say, for Normal Distribution one has mean=0 and sd=1).

Probability Distributions in R

Note that the distributions above have some parameters (say, min and max for Uniform, mean and sd for Normal etc.), and some parameters have default values you can keep (say, for Normal Distribution one has mean=0 and sd=1).

Plotting the Graph of Normal PDF and CDF

```
curve(dnorm, -4, 4, lwd = 2)
curve(pnorm, -4, 4, lwd = 2)

curve(dnorm(x, 1,3), -4, 4, lwd = 2)
```