

Multivariate Procedures with R

Prof. Shalabh

Department of Mathematics and Statistics

IIT Kanpur

Week – 01

Lecture – 02

Help, Demonstration, Examples, Packages and Libraries

Hello friends, welcome to the course multivariate procedure with R. So, in the last lecture you had seen that we had discussed about that how to get prepared by installing the R software and we had some brief information about the R software. Now, in this lecture we are moving forward. So, whenever you are trying to learn a new software what do you want? You always need some help. Sometime in a book when you want to understand a new topic the solved examples help you in understanding how you have to solve the problem and how you can understand the concept easily. So, the same story is in the R software also that whenever you are trying to learn new statistical procedure which are to be implemented in the R you would like to know that how you can do it.

Well, I am here only for some time and I am trying to cover here only a limited number of topics, but surely the journey is long and there will be many more topics which you would like to learn yourself. So, my objective here is that I will try to expose you with certain topics I have more confident on you that you can learn them yourself. So, whenever you are trying to learn a new statistical procedure which has to be implemented in the R software you always need some help, you always need some demonstration of those commands, you always need some examples to understand that how are you going to use them. So, in the R software there are different ways by which demonstration help examples of on a particular topic they can be obtained provided they are available.

So, now in this lecture I will try to give you a brief idea that how you can seek help and support in the R software to learn the new topics. Well, my lecture will be limited to only few topics, but these topics will make you understand that if you want to take the help of some other topic on the similar lines you can move forward right. So, let us begin our lecture. So, now in this lecture we are going to talk about the help, demonstration,

examples, packages and libraries. So, now you have installed the R software on your computer.

So, you will there have an icon like this one. So, you have to just double click on it and after that you will get here this type of window right. This is called as GUI that is Graphic User Interface G for graphic, U for user and I for interface and this is how this word GUI comes into existence. So, as soon as you start here you will see here there is some information written about the R software. We are not bothered about it and here you will see there are some standard buttons to execute these things.

So, this window is called here as a GUI window which is the common terminology. So, whenever we are talking of the GUI window that means you have to look for this window right. And now if I want to address that how you can seek the help in the R software then there are different ways right. First option is that if you come to this R GUI window as you can see here in this line there are different buttons right file, edit, view and similarly there is another button here help. So, if you come to help here then there is a console FAQ on R, FAQ on R for windows, manuals in PDF etcetera etcetera, HTML search help etcetera etcetera.

So, you can just click on any of this topic and then you can seek help in different formats. I will try to show you some examples here how are you going to do it. But this is just one basic approach and I always prefer if I can get the help on any topic by using the information which is available inside the R software. Reason being that whenever a new command or syntax is developed that is developed by some researcher, by some programmer and they are the people who are also going to create the help document. So, I believe that the person who has developed the syntax can guide us in the best way than anybody else.

So, I always try to depend on the help which is provided by the R software only. The second very common and popular option nowadays is to search over some internet site for example, you can use Google to the help. The reason being that now this R software has become very popular and there are several websites which are offering help. For example, many people are joining a platform and they try to address their issues, they try to raise their queries and some people try to address it, many times it helps and many time if you want to know that okay in order to this particular type of statistical analysis which of the package is going to be useful. So, under this case different people are trying to develop different types of packages and sometimes they try to host their details on their personal home pages also and they can update it and there are various types of

platforms which are trying to give an information that okay for the same job more than one packages can be used and these are the minor details.

So, those type of information many times helps us in understanding that what we want to do, right. And next approach is that suppose if you want to take help on any particular function. For example, suppose I want to know how to use the command matrix, M-A-T-R-I-X. Later on, we will use it and, in this course, actually you will see that whenever we are trying to deal with the multivariate procedures and then matrix theory offers a very good help in defining the data vectors and data matrix. So, that is why I have given this example here because anyway we are going to learn about the matrix theory in the R software, some basic commands.

So, if you want to know about this matrix, So, you have to just go to the GUI window and this is the GUI window and you can see here this was here you the prompt and here you have to type you simply type here question mark matrix. So, I am trying to give you only an example through this matrix, but if there is any other command you can write the question mark and then after that you write the command of the or the syntax whatever on which you want to take the help. Once you do it, it will do like this starting STTPD help server etc done and then after that it will come to this type of site, this site will be opened. And you can see here this is here the matrix which is written here inside the curly bracket is written here base that means it is inside the base package. What is base package?

So, the R software which you had installed earlier which you had downloaded from the rproject.org site that is called as the base package and as I told you that some of these packages are built in and some packages are contributed. So, this matrix package it is built in and it is available inside the base software. This is what it is trying to tell you and after that you can see here that it is trying to give you different command that matrix creates a matrix from the given set of values as dot matrix attempt to turn its argument into a matrix etc. And then here they are trying to give you the some commands and then they are trying to explain you here the meaning of the arguments for example, this argument they have to use inside the data for example, here this is here data, they are trying to explain you here that what do R expects in the data, right. And here is a command here something like nrow, ncol what does this mean, So, this nrow, ncol their interpretation, their meaning is explained here and So, on it will continue.

I will try to show you on the R console also but this is how you can get some information or at least the reliable information about the matrix command that how are you going to use it and you can see here this will continue. There are more details by row, dim names

and so, on and then it has all the minor details with their explanation and it is also given that from which book they have followed this command. So, let me try to first show you this thing on the R console. So, you can see here this is here the R console and this is here the R GUI window. Now and this point where I am highlighting this is here the command line, I will try to explain you but here you are going to type the commands.

So, if I try to type the command here, question mark, matrix you can and as soon as I say enter you can see here this page is open on your browser and you can see here that it is going to be like this I can see here. The same information which I shown you it is here, right. So, what I did that in order to explain you and in order to save some time I have taken the screenshot and you can compare here that this is the same screenshot which I had given you inside the lecture. So, yeah gradually you will follow this approach and then you will become more confident about the way we are going to handle this lecture, right. So, now I hope this clarifies that how you are going to use this, right.

Sometime you do not know the function or the command but you simply have some brief idea that what exactly you want to do, right. Suppose I want to do here the data input, suppose I want to know that how I can input the data. In such a case the one option is that if you simply type here `help.search`, H-E-L-P dot S-E-A-R-C-H and inside the parenthesis within the double quotes you try to write whatever you want data input for example I have typed.

Whenever you are trying to write down something within the double quote then it is taken as a comment, right, means as a character, right. So, if you try to use this `help.search` and if you try to do it in the R console it will happen like this that it will say starting HTTP help server and once I and then after that it will give me here this type of help screen. So, in this screen you can see here there will be several hyperlinks here and then you can see here that utilities like `read.table` or `read.df` they are going to provide you the help in reading the data from spreadsheet or for the data input. So, if you see here this data input is common in both this search outcome. So, now if you try to click over these links this one or this one then you will get more information, right. So, let me try to show you this information in the R console also. So, I will just copy and paste the same thing So, that I can save some time and if I come to here R console here if I try to see here Ctrl V you can see here it will start and it comes out to be here like this.

So, now if I try to click over this command `read.table` as you can see here it is about the data input and it is giving you the description that it reads a file in the table format and

creates a data frame from etc., etc. and then they have given here different types of options which can be used, right. So, this is what I wanted to show you.

Now let me come back to my lecture and then another option is that you can simply use here means another function which is here help and you simply say help and inside the parenthesis you do not write anything or you write help.start and then inside the parenthesis you do not write anything and it will open a HTML browser interface. For example, if you try to execute it on the R software it will come here like this. Now here you can see there are different option. This is about introduction, writing, R expression, import, export of the data about packages, right, and many many things are here.

So, now you can look into these details on various hyperlinks and can find out what exactly you want to do. Now after this another type of helps comes which can be obtained by two popular command or functions find, FIND and APROPOS, A-P-R-O-P-O-S and how they are going to help us. For example, about this find. Suppose you have this idea that you want to use some basic command to do some statistical analysis. Sometime it may happen that the command is available in the base package and sometime it may be available in some special contributed package.

So, later on I will explain you that when you want to use such contributed package first you have to upload it. But before that you have to know that this command is available in which of the package which has to be uploaded. So, in case if you want to find that this command is available in which of the package then you can use here the command like FIND. For example, if I want to know that where this matrix command is available in the base package or in some contributed package. So, if I simply say here FIND and within the double quote inside the parenthesis if I write M-A-T-R-I-X then it will return like this information package.

package colon base. This means this package is available in the base package, right. And you can see here this is how it will look like on the R software's GUI window. So, you can also execute it yourself and sometime you do not know the complete syntax but you have some idea. For example, I can recall that okay I had used earlier some command in which L-M was occurring. Means L-M were the two letters in the command and I cannot recall what was the exact command.

So, in such cases the command here APROPOS, A-P-R-O-P-O-S. If you try to use here and you try to write it the letters which you want to search for example here L-M inside

these double quotes then it will try to return you this type of detail and if you try to see what is happening. Now it is giving you here all the files in which or all the packages in which there is an L-M command. For example, you can see here it is L-M, right here it is L-M and if you try to see in every of this thing whatever is listed here this L-M is occurring.

For example, here G-L-M it has L-M. So, now by looking at these names you can possibly recall that which of the command you wanted to use. Well, I am not executing it on the R console because this is an outcome from my laptop where these commands are available which may be different than what your laptop or computer has. So, I do not want to create confusion So, that if you try to match this R software which the outcome of the R software which may not match with your laptop, So, I am just giving you here the screenshot, right. After that as we know that whenever we are trying to learn a new subject, new topic we always want to know about an example and if an example is present to us then the understanding becomes easy. So, similarly in the R software also there are various commands for which examples have been created and those examples can be accessed by using the command example, E-X-E-M-P-L-E and inside the parenthesis you have to write down the name of the function for which you need an example and in case it is available it will come.

For example, this L-M, L-M is a function which is used for the linear models, right. This is used in the linear models. So, there are different types of analysis which is done in the linear models that can be accessed through this L-M function. So, now I want to understand that how this L-M function will work. So, I will try to write example inside parenthesis L-M and I would try to execute it on the R console.

So, as soon as I do here, you will see here this type of screenshot will come, right. Say example L-M, you can see here required graphics, this, this, this and here the details about the book from which this example has been taken. So, it is giving you this is the, this book and the page number this, this data, all this data is there and after that they will try to give you the stepwise detail that how you are going to use it. So, if you want to use it, you have to just read it carefully and try to understand how it is going to do it here, right. So, before I try to go here, I will try to show you these things on the R console and in the R console if you want to clear the screen, the command here is control L.

So, if I try to press here control L, pressing the control button simultaneously with the L key, it will clear the screen. So, if I try to write down here example, see here L-M, you can see here it will come out to be here like this and if I try to confirm, enter it, different

types of graphics will come and so, on. So, you see those who are familiar with the regression analysis and linear models, they will be knowing that what it is trying to present here, right. So, let me try to clear the screen and come back to my lecture. Now once I have given you the example, then the next possibility comes that if somebody can give us a demonstration that how I am going to use it.

So, in R for some of the functions, these commands, this demonstration is available and this can be accessed by the command `demo`, D-E-M-O. So, for example, if I want to know about the perspective plots, these perspective plots are basically the three-dimensional surface plots and I want to know that how they will look like, how I can use it, etc. So, I can simply write down here `demo` and inside the parenthesis, `perp`, P-E-R-S-P, whatever is the correct spelling of this one and once I do it, you will see here this type of screen will come, it will try to give you the details that what type of data they have taken, how they are going to do it and then this type of graphics and if you try to press it more, it will try to give you more commands and similarly, if I want to know about the demonstration of the another useful command, say here `graphics`, So, I can use here `demo graphics` and in this case, you will see here this type of graphics will come here and it will give you here all the details that how it is going to be used here, right. So, let me try to first give you this example on the R software and then I will try to move to the next topic. So, if I try to give here demonstration of this perspective, you can see here if I say `enter`, this is giving me the details, now if I try to `enter`, it is giving me different types of graphics that how they can be created, right.

If you simply have to press `enter`, right and you can see here that how beautiful graphics can be created in the R software. The only thing is this, you have to see that what type of command, what type of functions you have to use in Kacha case and similarly, if you want to know about this graphics also, So, let me try to show you that how this will look like, So, let me clear the screen and close this window So, that I can start a new command for clarity, right. You can see here if I try to `enter`, I will be getting here different plots and examples with details which can be created with this graphics, right. Now the only thing is this, you can see here various types of possibilities are here for this creating this graphic. The only thing is this, you have to see what you want and how you can execute it.

So, let me clear the screen So, that we get prepared for the next command and now I try to show you here about using the libraries. So, now what do you mean by library? Can you recall in your school or college there is a library. What this library is doing? It has many books on different topics and whenever you want to learn any subject or if you want to understand any topic, you go to the library and you get the book issued, you bring

it to your home, office, hostel, you try to read it for some time, understand it and after that you give it back to the library. Now the next question is why cannot you have all the books at your home? Well, this is practically impossible. It is difficult to afford and then it is difficult to manage and we do not have that much space in our offices and home where we can keep all the books.

So, whenever I need it, I can go to the library and can buy the books. The same concept is being implemented in the R software also that whenever this R package was created, So, there were some options, some basic functionalities which are used by most of the people and there were some functionalities which were very specific to a specific type of users. So, all such common functionalities, they are compiled in base package and which is available on the home page of this R software and all other packages which are used by specialized people for some specialized jobs, they are compiled in different packages and which comes under the category of contributed packages. So, now if they try to combine all the packages, the base packages and contributed packages which different people would like to use, possibly the size of the software will become too much and everybody is not going to use it. So, definitely they would like to divide the packages into two component base and contributed.

So, base packages will always be there when you try to install the R software and contributed packages will be available on the website of the R software and whenever we want, we can download them, we can install them, we can use them and after that we can remove them So, that our space in the laptop or computer that is saved. So, this is the job which our library function in R software is doing. So, now our objective here is that we have to learn that how are we going to use the library in the R software and how are we going to install the packages, right. So, if you want to use any particular library, suppose you know the name of the library, then you have to use the command here library, libray, all in lowercase alphabets and then you have to write down library and inside the parenthesis you try to write down the name. For example, if I want to conduct the spatial analysis, right, spatial analysis is quite popular in statistics but it is not used by each and every user.

So, this is now the basic command, basic function which are used in the spatial analysis, they are combined in the library which is called as a spatial. Now I want to use it. So, whenever I want to use it, first I have to load the library. In order to load the library, I have to write down here library and inside the parenthesis spatial. Now the library is available in the R-GUI window for using it.

If it is not loaded, then if you want to use any command in this spatial analysis, it will not work. So, this is how we try to load it. There are certain packages which comes as a part of the base package. For example, there is a package here MASS mass, capital M capital A capital S. Actually, this is a package which is associated with the book modern applied statistics using S plus.

So, the first letter M, A, this S and this S, they are taken to create this package and to give it its name. Actually, this is the book which came along with the S plus software which was written by Van Bly, Van Bly and Ripley that was published by Springer and since this R package was developed on the lines of S plus software, So, all the datasets command etc. which were used in this book, they were compiled in this package mass and So, that people who were actually using S plus, they can easily migrate to the R software. They should be comfortable. So, this is now this mass package is available as a base package and similarly this MGCV, this is used for generalized additive model.

This is also available in the base package and many other commands are available in the base package, right. So, this is how we are going to use the library function to use any particular type of this package. Now whenever you are trying to use any library, you always want to know what is inside the library. So, for that you have a function here help. For example, if I say library help is equal to spatial that means I want to know what is inside this spatial library.

So, you will get here this type of information. All the details about this package, that package is spatial, priority recommended, version, date when it was created and all other things like here like this one. So, as I told you earlier that it is not a black box and all the information which is available in the R software that can be accessed by anyone, right. So, this is you will get it. Now the next question comes here, how are you going to bring these libraries from the R software? You do not have to go to the website of the R software but it can be obtained directly inside the R software on the R GUI window, right. So, as I said that the base R packages contains programs for some basic operations and it does not contain all the libraries for advanced statistical work.

So, specific jobs will have a specific requirement for that we are going to use some special packages. So, they are downloaded and downloading is very simple, right. To install any package you have to start the R software and on the R GUI window on the command line you simply have to write the command install.packages, right. And after that you have to write down the library which you want to install, right.

For example, I will just give you two examples. Suppose if you want to install the packages R meta which is used for the meta-analysis and the packages cluster which is used for the cluster analysis then you have to use the command like `install.packages` and within double course you write R meta and if you try to do it, it will look like this. As soon as you can see from this screenshot, as soon as you write here `install.packages R meta` within the double course, the first screen will come out to be like this where you have to choose your CRAN mirror from where you want to download it, right. And after that as soon as you choose it suppose if you want to choose Brazil or say any other country or any other server then you will get here this type of outcome.

So, finally you have to actually check that whether this package has been successfully installed or not. This you can check by looking at this line that package R meta successfully unpacked and MD5 some checked. So, that will ensure that the package has been downloaded and installed properly, right. Similarly, if you want to install the package cluster, So, you can see here `install.packages` and within parenthesis within double course you write here the cluster.

Cluster is a package which is used for the cluster analysis and if you try to do it here say `install.packages cluster` you will see here a similar type of say screen will come and finally you will get here a message like package cluster successfully unpacked and MD5 some checked. And yeah, this available in this library which is my actually my own laptop, right. But you may get a different message here.

But this message will remain the same. It is possible that some packages may become outdated with the new version of the R software but then it is not difficult to update them, right. And finally, if you want to quit from the R software you simply have to write down here `q` and this parenthesis. As soon as you do you will come out of the R software. For example, if I try to show you here that how it is going to be done, right. So, if I come to the R window and if I say here RGI window, if I say here RQ you can say here it will ask you save workspace image.

Now whatever work you have done here if you want to save then you say yes. If you do not want to save then you say no. And if you do not want to do anything, just say cancel. Suppose I want to quit and I want to save my workspace that means whatever I have done up to now this should be saved. I will say here yes and you can see here then the R is now terminated and we are out of the this RGI window. So, now I would like to stop in this

lecture here but if you try to see I have given you a brief idea about the working, the way we are going to seek help from the R software So, that our further journey is easier.

Especially when we are trying to learn new topics, new subject we always need help. In my personal opinion without help even it is the life also becomes very difficult. So, it is a better option to seek the help from your family, from your friends So, that our life becomes beautiful and similarly if you try to seek the help in learning the new topics our life also becomes easier. And the only thing is this we have to seek the help from the right people and similarly in the R software also we need to seek the help from the right source. In case if you try to use some wrong sources, they will possibly give you a wrong advice, wrong information and the life will become difficult but that is just like a simple philosophy of life.

So, you try to practice these things, try to see how you can do it and then try to brush up your knowledge and I will see you in the next lecture with more details on the R software till then goodbye.