



دانشگاه اراک

۱۴۰۰

# محاسبات آماری با R

کتاب دکتر محمدمرادی و عبدالله جلیلیان

ارائه دهنده:

محراب عتیقی

صفحات: ۱ - ۲۰

- نرم افزار R یک نرم افزار محاسباتی آماری است که در زمان کوتاهی مورد استقبال گسترده قرار گرفته است. با اطمینان می توان گفت که R هر روز از روز پیش کامل تر است، زیرا بیشتر کاربران آن با نوشتن برنامه هایی برای حل مسأله های خود و به اشتراک گذاشتن آنها در پیشرفت R نقش دارند. این نرم افزار یک زبان برنامه نویسی و محیطی یک پارچه را در اختیار کاربران قرار می دهد تا به کمک آن بتوانند کارهای گوناگونی از جمله موارد زیر را انجام دهند:
- وارد کردن، ویرایش، پیکره بندی و ذخیره ی انواع داده ها در قالب های مناسب،
- اجرای روش ها و تحلیل های رایج آماری با به کار بستن تابع های موجود در R،

- اجرای روش ها و تحلیل های جدید آماری با به کارگیری بسته های نرم افزاری موجود بر پایگاه های نرم افزاری مانند CRAN, GitHub و BioConductor،
- اجرای محاسبات عددی مانند اعمال جبری روی بردارها، ماتریس ها و آرایه ها، حل معادله ها خطی و ناخطی، یافتن ریشه ی چند جمله ای ها، بهینه سازی، تقریب انتگرال، مشتق گیری و...
- رسم انواع نمودارهای آماری با انعطاف پذیری بالا،
- به کار بستن یک زبان برنامه نویسی ساده و کارا برای پیاده سازی الگوریتم ها و هر نوع تحلیل آماری و ریاضی دلخواه.

در واقع محیط  $R$  علاوه بر تحلیل داده‌ها و رسم نمودارهای گرافیکی، بستری مناسب فراهم می‌کند که می‌توان به کمک زبان برنامه نویسی  $R$ ، (مبتنی بر زبان برنامه نویسی  $S$ ) به طور تقریبی هر نوع برنامه‌ای برای انجام محاسبات آماری دلخواه را می‌توان نوشت. به عنوان نمونه، ممکن است کاربر به محاسبه‌ی گونه‌ی جدیدی از آماره‌ی آزمون بر اساس یک روش تکراری، نمودار آماری ویژه‌ای با ساختار دلخواه و یا مطالعه‌ی شبیه‌سازی خاصی نیاز داشته باشد که دستورهای مربوط به آن در تابع‌ها و بسته‌های نرم‌افزاری موجود، یافت نشود.

در این صورت وی می‌تواند با به کار بستن قابلیت های برنامه نویسی R برنامه ای برای انجام محاسبه یا رسم نمودار مورد نظر خود بنویسد.

در این فصل با بیان تاریخچه‌ای از پیدایش R، به روش دریافت، نصب و آشنایی با محیط آن می‌پردازیم و در فصل های بعد روش وارد کردن داده‌ها، ویرایش داده‌ها، رسم نمودارها و تحلیل های آماری داده‌ها را توضیح می‌دهیم.

در اواسط قرن بیستم که انواع روش‌های آماری برای تحلیل داده‌ها توسط آماردانان و پژوهشگران دیگر رشته‌ها معرفی و در سطح گسترده‌ای از علوم و رشته‌های مهندسی فراگیر شده بودند، برای اجرای محاسبات آماری، بیشتر زبان‌های برنامه‌نویسی رایج آن دوران، به ویژه فورترن، به کار گرفته می‌شد. به همین خاطر انجام محاسبات مربوط به تحلیل داده‌ها به نوشتن برنامه‌های طولانی و مهارت بالای برنامه‌نویسی نیاز داشت. در سال‌های ۱۹۷۵ - ۱۹۷۶ به منظور انجام ساده‌تر و سریع‌تر محاسبات آماری و عددی در مجموعه‌ی مشهور و با سابقه‌ی آزمایشگاه‌های بل نسخه‌های آغازین زبان برنامه‌نویسی S توسط جان چمبرز و با همکاری ریک بکر و آلن ویک شکل گرفت.

در سال ۱۹۸۸ تغییرات زیادی در زبان S ایجاد و ویژگی‌های جدیدی مانند به کارگیری تابع نویسی به آن افزوده شد و نسخه‌ی جدید "new S" نام گرفت. سپس ویژگی‌های برنامه نویسی شی‌گرا مانند به کارگیری کلاس‌ها و متدها نیز به S افزوده و نسخه‌ی جدید S4 نامیده شد.

در این میان به کارگیری زبان S محدود به آزمایشگاه‌های بل نشد و این زبان در اختیار دیگر افراد به ویژه پژوهشگران در دانشگاه‌ها و مراکز علمی قرار گرفت. سادگی برنامه نویسی با S، امکان اجرای روش‌های آماری رایج با دستورهای ساده و کوتاه، رسم انواع نمودارهای گرافیکی و انتشار چند کتاب در مورد شیوهی به کارگیری S توسط تیم توسعه دهندهی آن باعث شد این زبان مورد توجه قرار گیرد و گسترش یابد.



به دلیل استقبال از S، یک پیاده سازی تجاری از آن با عنوان S\_PLUS توسط یک شرکت تجاری منتشر و برای فروش به بازار عرضه شد. در سال ۱۹۹۳ راس ایهاکا و رابرت جنتلمن از دانشگاه اوکلند نیوزیلند با افزودن ویژگی هایی از زبان برنامه نویسی Scheme به S، پروژهی توسعهی یک پیاده سازی غیر تجاری و متن باز از S را آغاز کردند و آن را R نامیدند: حرف نخست نام هر دو نفر و حرف قبل از S در الفبای انگلیسی. پس از آن متخصصان دیگری از سراسر جهان، از جمله جان چمبرز خالق S، به آنها پیوستند و تیم هسته‌ی توسعهی R را تشکیل دادند که تا به امروز مسئولیت توسعه و به روز رسانی R را به عهده دارد.

## ۲.۱ دریافت و نصب نرم افزار

نرم افزار R رایگان و تحت پروانه‌ی عمومی استفاده‌ی همگانی گنو (GNU) از پایگاه اینترنتی پروژه‌ی R به آدرس [www.r-project.org](http://www.r-project.org) قابل دریافت است و به سادگی نصب می‌شود. شبکه‌ی جامع بایگانی R با عنوان اختصاری CRAN، از مجموعه‌ای از سرورهای دانشگاه‌ها و مراکزهای پژوهشی در سراسر جهان تشکیل شده است و نسخه‌ی یکسان و به روز را در اختیار همگان قرار می‌دهد. در ایران دانشگاه فردوسی مشهد با سرور <http://cran.um.ac.ir> عضوی از این شبکه است. برای دریافت R می‌توان از این سرور استفاده کرد و یا با مراجعه به پایگاه اینترنتی [www.r-project.org](http://www.r-project.org) و انتخاب گزینه‌ی CRAN از منوی سمت چپ، از فهرست سرورهای موجود در سراسر جهان یکی را انتخاب کرد.

## ۲.۱ دریافت و نصب نرم افزار

11

سایت R پیشنهاد کرده است سروری که به شما نزدیک است را انتخاب کنید. سپس بر اساس سیستم عامل کامپیوتری که برنامه را روی آن نصب میکنید، نسخه‌ی مناسب برای سیستم عامل خود را انتخاب کنید. نرم افزار R برای سیستم عامل های لینوکس، مک و ویندوز موجود است. با توجه به رایج تر بودن سیستم عامل ویندوز، در ادامه فرض می‌شود گزینه‌ی **Download R for Windows** توسط کاربر انتخاب شده است. در صفحه‌ای که باز شده است با کلیک بر روی عبارت **Install R for the first time** به صفحه‌ی جدیدی منتقل می‌شوید که با کلیک بر روی عبارت **Download R 3.4.1 for Windows** آخرین نسخه (R) در حال حاضر نسخه‌ی ۴.۱.۳، اما با توجه به به‌روز رسانی پیوسته‌ی R ممکن است شماره‌ی نسخه برای شما متفاوت باشد دریافت می‌شود. پس از دریافت R، می‌توان آن را با دنبال کردن گام های زیر بر روی سیستم عامل ویندوز نصب کرد.

1. بعد از دانلود، به مسیری در کامپیوتر شما که فایل دریافت شده و با نام **R-4.1.3-win.exe** در آن ذخیره شده است، مراجعه کنید و با کلیک کردن بر روی فایل آن را اجرا کنید.
2. در پنجره ای که باز می شود، زبان مورد استفاده در گام های نصب (به صورت پیش فرض انگلیسی) را انتخاب و روی گزینه ی **OK** کلیک کنید.
3. در پنجره ی جدید روی گزینه ی **Next** کلیک کنید.
4. در پنجره ای که باز می شود، پروانه ی گنو را مطالعه و در صورت موافق بودن با شرایط آن روی گزینه ی **Next** کلیک کنید.

## ۲.۱ دریافت و نصب نرم افزار

5. در پنجره‌ی بعدی مسیری را که می‌خواهید در آن نصب شود، با کلیک بر روی گزینه‌ی **Browse** تعیین و یا مسیرپیش‌فرض (C:\ProgramFiles\R) را نگه دارید و روی گزینه‌ی **Next** کلیک کنید.
6. در این گام مؤلفه‌هایی که باید روی سیستم عامل نصب شوند مانند فایل‌های هسته (**Core Files**)، فایل‌های ۳۲ بیتی، فایل‌های ۶۴ بیتی و ترجمه‌ی پیغام‌ها را می‌توان انتخاب کرد. به صورت پیش‌فرض همه‌ی مؤلفه‌ها انتخاب شده‌اند. برای ادامه روی گزینه‌ی **Next** کلیک کنید.
7. در این پنجره از شما پرسیده می‌شود که آیا مایل به تغییر شیوه‌ی راه اندازی R هستید؟ به صورت پیش‌فرض گزینه‌ی **No (accept defaults)** انتخاب شده است. این گزینه را تغییر ندهید مگر آن که با R و بسته‌های آن آشنایی کافی داشته باشید. روی گزینه‌ی **Next** کلیک کنید.

8. در این پنجره به شما اطلاع داده می شود که یک میانبر از برنامه در منوی استارت با نام R (حالت پیش فرض) ایجاد خواهد شد. روی گزینه ی **Next** کلیک کنید.
9. در پنجره ی جدید نیز به شما اطلاع داده می شود که آیکون برنامه بر روی میزکار ویندوز قرار خواهد گرفت، اطلاعات مربوط به نسخه ی برنامه در **Registry** ویندوز ثبت خواهد شد و فایل های با پسوند **Rdata**. توسط R باز خواهند شد. برای ادامه روی گزینه ی **Next** کلیک کنید. فرایند نصب آغاز می شود و می توانید میزان پیشرفت آن را ببینید.
10. در صورتی که نصب برنامه موفقیت آمیز باشد، پنجره ی باز می شود که به شما اطلاع می دهد نصب R با موفقیت انجام شده است. برای تمام فرآیند نصب بروی گزینه ی **finished** کلیک کنید.

Download and Install R and RStudio | R Tutor

725 observations of 6 variables

	LungCap	Age	Height	Smoke	Gender	Caesarean
1	6.475	6	62.1	no	male	no
2	10.125	18	74.7	yes	female	no
3	9.558	10	69.7	no	female	yes
4	11.125	14	71.8	no	male	no
5	6.889	5	56.9	no	male	no
6	6.225	11	58.7	no	female	no
7	4.958	8	63.3	no	male	yes

Console

R is a collaborative project with many contributors. Type 'contributors()' on the command line for more information and to see who is currently contributing. Type 'demo()' for some demos, 'help()' for on-line help, or 'help.start()' for an HTML browser interface to help. Type 'q()' to quit R.

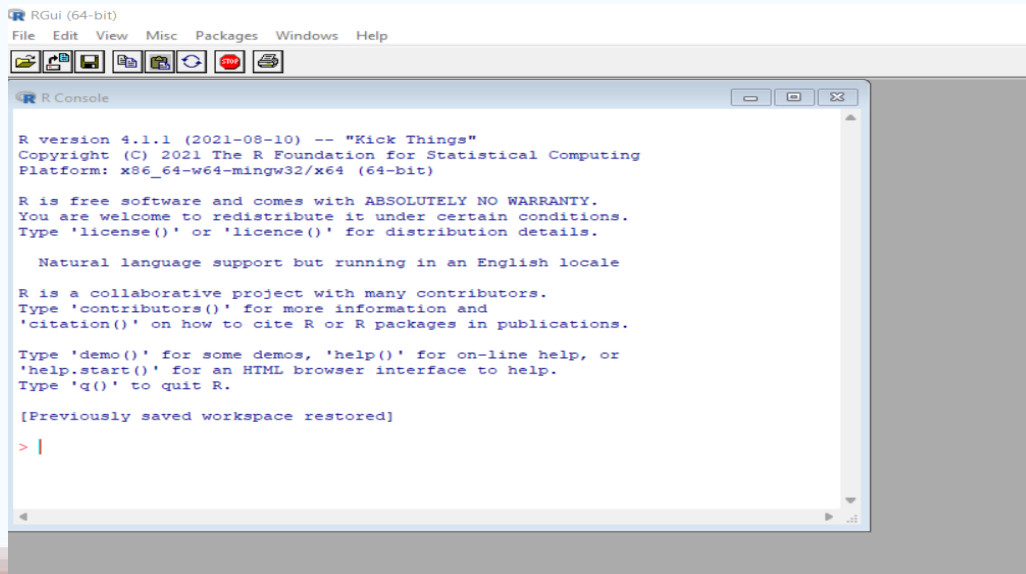
```
> LungCapData <- read.delim("~/Desktop/R Videos Project May 2013/LungCapData.txt")
> View(LungCapData)
> attach(LungCapData)
> names(LungCapData)
[1] "LungCap" "Age" "Height" "Smoke" "Gender" "Caesarean"
> plot(Age, lungCap, main="Lung Capacity by Age")
```

Mike Marin  
Ladan Hamadani

Lung Capacity by Age

12:43 PM

چنانچه نرم افزار R با موفقیت بر روی کامپیوتر نصب شده باشد، با کلیک کردن روی میان‌بر آن روی میز کار و یا منوی استارت ویندوز، پنجره‌ی رابط گرافیکی R با نام Rgui باز شده و مانند شکل زیر محیط R در دسترس قرار خواهد گرفت.





پنجره‌ی اصلی رابط گرافیکی R از تعدادی منو، تعدادی میان‌بُرد در نوار ابزارها و یک پنجره با عنوان R Console تشکیل شده‌است. منوهای File , Edit , View , Windows , Help کارکردی همانند بیش‌تر نرم‌افزارهای تحت سیستم عامل ویندوز دارند و منوهای Misc , Packages به R و بسته‌های نرم‌افزاری می‌پردازند که به R افزوده می‌شوند. آیکون‌های موجود در نوار ابزارها نیز میان‌بُرهایی برای برخی از گزینه‌های داخل منوها هستند.

اصلی ترین بخش در محیط R پنجره‌های R Console است که رابط بین کاربر و R خواهد بود. در این پنجره یک خط فرمان (باعلامت >) وجود دارد که مکان‌نمای جلوی آن آماده‌ی دریافت دستورها از کاربر است.

&gt;

برخلاف نرم‌افزارهایی مانند SPSS , Minitab که تعامل کاربر و نرم‌افزار با کلیک کردن بر روی منوهای از پیش تعیین شده است، ارتباط کاربر با R، به کمک این خط فرمان امکان پذیر است. همه‌ی کارها از جمله ورودی و ویرایش داده‌ها تا انجام انواع تحلیل‌های آماری، با وارد کردن دستور مناسب توسط کاربر و اجرای دستور، توسط R صورت می‌گیرد.

کاربر با تایپ دستورها در خط فرمان و فشردن کلید **Enter** روی صفحه‌ی کلید اجرای محاسبات مورد نظر خود را از **R** درخواست می‌کند. در صورتی که دستورها به درستی نوشته شوند، **R** این دستورها را اجرا می‌کند که ممکن است به چاپ خروجی در همین پنجره، رسم نمودار در پنجره‌ی گرافیکی و یا موردهای دگر بیانجامد. در مواردی هم مانند ذخیره‌کردن یک ساختار داده‌ای در یک شیء بعد از اجرای دستور، مکان‌نما به خط بعد منتقل می‌شود و هیچ پیامی مبنی بر خطا یا خروجی به‌دست آمده از دستور دیده نمی‌شود.

برای مثال، اگر دستور `R.version` را اجرا نماییم، اطلاعاتی درمورد سیستم عامل و نوع معماری کامپیوتر ما و نسخه‌ای از R که بر روی آن نصب است به‌عنوان خروجی چاپ می‌کند.

```
> R.version
```

```
platform      _x86_64-w64-mingw32
arch           x86_64
os             mingw32
system        x86_64, mingw32
status
major          4
minor          1.1
year           2021
month          08
day            10
svn rev        80725
language       R
version.string  R version 4.1.1 (2021-08-10)
nickname       Kick Things
```

نکاتی که در زیر بصورت یک جدول آمده است از مباحث پایه ای و مقدماتی R می باشند.

ردیف	مثال	شرح	تابع
۱	ls()	فهرست نام شیء های ایجاد شده	Ls
۲	objects()	فهرست نام شیء های ایجاد شده	Objects
۳	data()	فهرست داده ها موجود	Data
۴	rm()	حذف یک شیء مانند X	rm
۵	save(x , y , file = "my.Rda")	ذخیره کردن شیء ها در یک فایل	save
۶	print(x)	چاپ یک شیء مانند X	print
۷	cat("t = " , x)	چاپ و پیوست یک شیء در خروجی مانند کارکترهای t = و مقدار x	cat
۸	help(mean)	نمایش صفحه ی راهنمای مربوط به یک شیء مانند تابع mean	help
۹	help.search("geometric")	جستجوی صفحه های راهنما، برای یک عبارت مانند geometric	help.search

# ۱.۳.۱ نوشتن دستورها در یک فایل اسکریپت

در برخی موارد ممکن است نیاز به نوشتن یک برنامه باشد؛ یعنی باید چندین خط از دستورها نوشته شوند و سپس اجرا شوند تا نتیجه‌های مورد نظر به دست آیند. در چنین شرایطی بهتر است دستورها در یک فایل اسکریپت وارد شوند. برای ایجاد یک فایل اسکریپت جدید، از منوی **File** گزینه‌ی **New script** را انتخاب کنید. در این صورت پنجره‌ی یک ویرایشگر متن در محیط **R** باز می‌شود و دستورها را می‌توان در این پنجره تایپ کرد.

پس از ورود دستورهای مورد نظر، برای اجرا می‌توان نخست بخش مورد نظر از دستورهای نوشته شده را انتخاب و سپس کلیدهای **Ctrl** و **R** را هم‌زمان فشرد تا دستورها به پنجره‌ی **R Console** کپی و سپس اجرا شوند. پس از انتخاب دستورها، می‌توان به جای فشردن **Ctrl** و **R** بر روی آیکن **Run line or selection** در نوار ابزار کلیک کرد. با به کارگیری گزینه‌ی **Save** از منوی **File** و با کلیک بر روی آیکن مربوط به ذخیره سازی در نوار ابزار می‌توان فایل اسکریپت ایجاد شده را برای استفاده‌های بعدی به کار برد. فایل‌های اسکریپت با پسوند **R**. ذخیره می‌شوند و برای باز کردن آنها در محیط **R** می‌توان از منوی **File** گزینه‌ی **Open script** را انتخاب و یا روی آیکن مربوط به آن در نوار ابزار کلیک کرد.

برای ذخیره‌ی نمودارهای رسم شده در **R** نیز می‌توان از منوی **File** با کلیک بر گزینه‌ی **Save as** نمودار را انتخاب و به یکی از قالب‌های گرافیکی رایج **Jpg**، **TIFF Bmp**، **Png**، **WMF**، **PDF** و **Postscript** ذخیره کرد.

# ۱.۳.۱ نوشتن دستورها در یک فایل اسکریپت

برای به کارگیری نمودار ذخیره شده در نرم افزارهای حروفچینی مانند Microsoft Word قالب WMF و برای به کارگیری نمودار در سیستم های حروفچینی مبتنی بر TEX مانند DTEX و Xpersian قالب های PostScript و PDF مناسب هستند. رسم انواع نمودارها در در فصل ۳ شرح داده می شود. برای ذخیره ی مطالب موجود در پنجره ی R Console که دستورهای اجرا شده و خروجی های آنها را در بر دارد، نخست با کلیک بر روی آن، پنجره را فعال و سپس از منوی File با کلیک بر گزینه Save to File محتوای پنجره ذخیره می شود. ذخیره، باز کردن و بارگذاری قاب ها به صورت پیش فرض در دایرکتوری کاری ( Working dicrectory ) R انجام می شود. با دستور

```
> getwd()
> [1] "C:/Users/Frostless/Documents/R"
```

می توان مسیر کامل این دایرکتوری را دید و با دستور

```
> Setwd("E:/Project")
```

می توان آن را به مسیر دلخواه خود ( برای نمونه، دایرکتوری MyProject در درایو E کامپیوتر) تغییر داد دقت کنید که برخلاف مسیرهای رایج در ویندوز، در تعیین مسیر فایل ها و دایرکتوری ها در R علامت / به جای \ برای جدا کردن زیردایرکتوری ها به کار برده می شود.

علاوه بر دستورها و بسته‌های نرم‌افزاری پایه‌ی R می‌توان هزاران بسته‌ی نرم‌افزاری ویژه‌ی R را به کار برد که افراد مختلف از سراسر دنیا برای مقاصد گوناگون آن‌ها را تهیه و به صورت رایگان و متن باز در اختیار همگان قرار داده می‌دهند. همچنین می‌توان بسته‌های نرم‌افزاری را تهیه و به کتابخانه‌ی R افزود. رایگان بودن و مشارکت عمومی آمادان‌ها و کاربران R از سراسر دنیا باعث شده که در چند سال اخیر این نرم‌افزار به نرم‌افزاری کامل و فراگیر تبدیل شود. برای دیدن فهرست بسته‌های نرم‌افزاری که برای R تهیه شده و از CRAN در دسترس هستند می‌توان با مراجعه به پایگاه اینترنتی

[www.r-project.org](http://www.r-project.org) و انتخاب گزینه‌ی CRAN از منوی سمت چپ صفحه، از فهرست سرورهای موجود در سراسر جهان یکی را انتخاب کرد. سپس با کلیک بر روی Packages در منوی سمت چپ صفحه، فهرست بسته‌های نرم‌افزاری به ترتیب تاریخ انتشار یا به ترتیب نام (حروف الفبا) قابل مشاهده است.



. در صورت نیاز می‌توان هر یک از این بسته‌های نرم‌افزاری را در قالب یک فایل فشرده (Zip) دریافت و سپس به R افزود. برای این منظور کافی است پس از دریافت فایل فشرده‌ی بسته‌ی نرم‌افزاری مورد نظر، در پنجره‌ی رابط گرافیکی R از منوی **Packages** گزینه‌ی **Install package(s) from local zip files** را انتخاب و در پنجره‌ای که باز می‌شود به مسیری که بسته نرم‌افزاری در آن ذخیره شده رفت و آن را انتخاب چنانچه کامپیوتر به اینترنت متصل است، می‌توان بسته نرم‌افزاری مورد نظر را به طور مستقیم از محیط R دریافت و نصب کرد. برای این منظور، با انتخاب گزینه‌ی **install package(s)** از منوی **Packages** پنجره‌ای باز خواهد شد که باید یکی از سرورهای فهرست شده در آن را انتخاب کرد. پس از انتخاب سرور، پنجره‌ی فهرست همه‌ی بسته‌های نرم‌افزاری موجود بر **CRAN** به ترتیب نام (حروف الفبا) ظاهر می‌شود که می‌توان بسته‌ی نرم‌افزاری مورد نظر را انتخاب کرد. در این صورت بسته‌ی نرم‌افزاری انتخاب شده دریافت و سپس نصب خواهد شد. در این حالت به جای به کارگیری منوها، می‌توان دستور

```
➤ Install.packages("spatstat")
```

را برای دریافت و نصب بسته‌ی نرم‌افزاری مورد نظر در این جا برای نمونه **spatstat** به کار برد.

پس از دریافت و نصب هر بسته نرم‌افزاری، آن بسته به مجموعه‌ی بسته‌های پایه‌ای R نصب شده روی کامپیوتر افزوده می‌شود و از آن پس می‌توان آن را به کار برد. دستور زیر برای فراخوانی بسته‌ی نصب شده می‌باشد. البته از منوی **Packages** گزینه‌ی **Load Packages** نیز می‌توان این کار را انجام داد.

```
➤ library(spatstat)
```

اگر این دستور را بدون اینکه هیچ بسته‌ای درون آن بنویسیم، آنگاه فهرست تمامی بسته‌های موجود و نصب شده روی کامپیوتر شمارا نمایش می‌دهد. دستور زیر نیز برای دیدن مسیر نصب بسته‌های نصب شده در سیستم داریم:

```
➤ .libPaths()
```

```
> 3 + 3  
[1] 6  
> 3 - 3  
[1] 0  
> 3 * 3  
[1] 9  
> 9 / 3  
[1] 3  
> 9 / 6  
[1] 1.5  
> 3 ^ 2  
[1] 9  
> 9 %% 6  
[1] 3  
> 9 %/% 6  
[1] 1
```

```
> abs( -5.7 )  
[1] 5.7  
> sign(-5.7)  
[1] -1  
> sqrt( 9 )  
[1] 3  
> exp ( 1 )  
[1] 2.718282  
> log ( 5 )  
[1] 1.609438  
> log10 (10)  
[1] 1  
> log2 (4)  
[1] 2  
> floor(1.3)  
[1] 1  
> round(4.6579)  
[1] 5
```

```
> sin(1.5)
[1] 0.997495
> asin(0.997495)
[1] 1.5
> sinh(1.5)
[1] 2.129279
> asinh(2.129279)
[1] 1.5
> factorial(5)
[1] 120
> choose(5 , 3)
[1] 10
> gamma(1.5)
[1] 0.8862269
> beta(1.5 , 5)
[1] 0.07388167
```

# ۴.۱ انجام محاسبات ریاضی در R

ردیف	دستور	کاربرد
۱	+	عملگر جمع
۲	-	عملگر تفریق
۳	*	عملگر ضرب
۴	/	عملگر تقسیم
۵	^	عملگر توان رسانی
۶	%/	عملگر باقی مانده تقسیم
۷	%/%	عملگر خارج قسمت تقسیم

ردیف	دستور	کاربرد
۱	abs	قدر مطلق
۲	sign	تابع علامت
۳	sqrt	جذر
۴	exp	تابع نمایی بر پایه ی $e$
۵	log	لگاریتم بر پایه ی عدد طبیعی $\ln$
۶	log10	لگاریتم بر پایه ی عدد ۱۰
۷	log2	لگاریتم بر پایه ی عدد ۲
۸	floor	جزء صحیح

ردیف	دستور	کاربرد
۱	round	گرد کردن تا چند رقم اعشار
۲	sin	سینوس
۳	asin	سینویس اینورس
۴	sinh	سینوس هایپربولیک
۵	asinh	سینویس هایپربولیک اینورس
۶	factorial	فاکتوریل
۷	choose	تعداد ترکیب ها
۸	gamma	تابع گاما
9	beta	تابع بتا



برای انجام عمل مشتق‌گیری در R داریم،

```
> D(expression(3 * sin(x^5)) , "x" )  
3 * (cos(x^5) * (5 * x^4))  
> D(expression(3 * x^5) , "x" )  
3 * (5 * x^4)  
> D(expression(log(3 * x^5)) , "x" )  
3 * (5 * x^4)/(3 * x^5)
```

## ۲.۴.۱ اعداد مختلط و مقدارهای عددی ویژه در R 34

علاوه بر اعداد حقیقی (صحیح و اعشاری) تعریف اعداد مختلط و انجام اعمال حسابی مربوط به آنها در R نیز امکان پذیر است. به عنوان نمونه، عدد مختلط  $1 + 3i$  در R به صورت زیر تعریف و محاسباتی و توابع ریاضی به کار گرفته می شود.

```
> 1 + 3i  
[1] 1+3i  
> 1 + 3i + 5.5 - 5i  
[1] 6.5-2i  
> (1 + 3i)^3  
[1] -26-18i  
> exp(1 + 3i )  
[1] -2.691079+0.383604i
```

## ۲.۴.۱ اعداد مختلط و مقدارهای عددی ویژه در R

ما در زبان برنامه نویسی R به مقادیری ویژه نیاز داریم و می‌بایست آشنایی اولیه‌ای با آنها داشته باشیم و لذا یک لیستی از آنها تهیه کردیم در پایین.

ردیف	دستور	شرح
1	NaN	مقدار عددی تعریف نشده است
2	NA	مقدار گم‌شده است
3	-Inf , Inf	مثبت و منفی بینهایت
4	NULL	بوج
5	pi	عدد $\pi$
6	TRUE	مقدار منطقی درست
7	FALSE	مقدار منطقی نادرست

# ۳.۴.۱ دقت محاسبات عددی در کامپیوتر

در مورد محدودیت تعداد رقم‌های معنی‌دار یک عدد و محاسبات عددی در کامپیوتر پاره‌ای توضیحات لازم است. به عنوان نمونه عدد گنگ

$$\pi = 3.14159265$$

و عدد گویای

$$\frac{2}{3} = 0.66666666$$

را در نظر بگیرید. چنین عددهایی دارای بسط اعشاری متناهی نیستند؛ یعنی تعداد رقم‌های اعشاری آن‌ها بی‌نهایت است. از آنجایی که عددها به صورت رشته‌هایی با طول متناهی از صفرها و یک‌ها در حافظه‌ی کامپیوتر قرار می‌گیرند، چنین عددهای حقیقی را نمی‌توان در کامپیوتر ذخیره کرد و در عمل تنها مقدار تقریبی آنها (با تعداد محدود رقم اعشاری) به کار برده می‌شود. همچنین ذخیره کردن اعداد بسیار بزرگ مانند

$$3 * 10^{500}$$

و یا اعداد بسیار کوچک مانند

$$2.2 * 10^{-452}$$

با چالش‌هایی همراه است و به دلیل متناهی بودن حافظه هر کامپیوتری امکان ذخیره کردن هر عدد بسیار بزرگ یا بسیار کوچک امکان‌پذیر نیست. بنابراین هر عدد حقیقی دلخواهی را نمی‌توان در کامپیوتر ذخیره کرد. در واقع اگر در یک کامپیوتر نوعی هر عدد را بتوان در یک رشته به طول  $N$  ذخیره کرد تنها  $2^N$  عدد صحیح و تنها  $2^N$  اعداد اعشاری در بازه  $(0,1)$  را می‌توان آرایه داد. این در حالی است که مجموعه‌ی اعداد صحیح نامتناهی شمارا و مجموعه اعداد حقیقی در بازه  $(0,1)$  نامتناهی ناشمار است.

# ۳.۴.۱ دقت محاسبات عددی در کامپیوتر

برای به کارگیری بهینه ی حافظه و امکان ذخیره کردن گستره بیشتری از عددهای بزرگ و کوچک، اعداد در کامپیوتر با نمایش ممیز شناور (نماد علمی به شکل

$$x = a * 10^c$$

ارایه می شوند که در آن عدد حقیقی  $0 \leq a < 10$  و ضریب علمی، ۱۰ پایه و عدد صحیح  $c$  توان است. به عنوان نمونه، عدد  $x = ۱۵۲۸۵۳.۵۰۴۷$

دارای نمایش ممیز شناور  $x = 1.528535047 * 10^5$  است. اگر یک

کامپیوتر تنها بتواند عددهای با ضریب علمی  $a$  تا ۷ رقم معنی دار (صحیح و اعشاری) را ذخیره کند، آن گاه  $x$ ، به صورت  $1.528535 * 10^5$  (رقم های ۰۴۷ از انتهای ارقام اعشاری  $x$  بریده شده است) ذخیره می شود که با مقدار واقعی آن اندکی اختلاف دارد. علاوه بر این، نتایج عمل های حسابی مانند جمع در کامپیوتر ممکن است با آنچه ما همه جا از اعداد حقیقی انتظار داریم اندکی متفاوت باشد.

## ۳.۴.۱ دقت محاسبات عددی در کامپیوتر

به عنوان نمونه، برای این که دو عدد با ممیز شناور را با هم جمع کنیم نخست باید هر دو عدد با توان یکسانی نمایش داده شوند و سپس جمع شوند. پس اگر بخواهیم  $123456.7$  را با  $101.7654$  جمع بکنیم، داریم:

$$\begin{aligned} & 123456.7 + 101.7654 \\ &= 123456.7 + 0.001017654 \times 10^5 \\ &= (1.234567 + 0.001017654) \times 10^5 \\ &= 1.235584654 \times 10^5 \end{aligned}$$

بنابراین در محاسبه با کامپیوتری که بتواند عددهای با ضریب علمی تا  $10^5$  رقم معنی دار را ارایه دهد، نتیجه‌ی جمع به صورت

$123558.5 = 1.235585 \times 10^5$  نمایش داده خواهد شد. پس عمل‌های حسابی با عددهای با ممیز شناور ممکن است با آنچه ما انتظار داشته باشیم اندکی اختلاف داشته باشند.

## ۳.۴.۱ دقت محاسبات عددی در کامپیوتر

نکته‌ی دیگری که درمورد محاسبه‌های عددی در کامپیوتر باید به آن توجه داشت این است که عددها در حافظه کامپیوتر به صورت دودویی (درپایه‌ی ۲) ذخیره و عملگرهای محاسباتی (جمع و جایگزینی) نیز بصورت دودویی انجام می‌شوند. این امر باعث می‌شود که نتیجه‌ی برخی محاسبات که به صورت دودویی انجام می‌شوند، با آنچه ما در پایه‌ی ۱۰ انتظار داریم اندکی اختلاف داشته باشند. به مثال زیر توجه نمایید:

```
> 1.12e-7 +1  
[1] 1  
> 0.5 - 0.4 - 0.1  
[1] -2.775558e-17  
> sin(pi)  
[1] 1.224606e-16
```

## ۴.۴.۱ عبارتهای مقایسه‌ای و منطقی در R

گاهی نیاز است مقادارهای گوناگون از نظر برابری، کوچک‌تر یا بزرگ‌تر بودن مقایسه شوند. عملگرهای مقایسه‌ای در R در زیر آورده شده‌اند:

ردیف	عملگر	منطق و شرح
۱	$<, >$	بزرگ‌تر و کوچک‌تر
۲	$<=, >=$	بزرگ‌تر مساوی و کوچک‌تر مساوی
۳	$==$	مساوی
۴	$!=$	متضاد
۵	$\&\&, \&$	و منطقی
۶	$  ,  $	یا منطقی
۷	$!$	نقیض



محاسبات مربوط به توزیع‌های احتمال پرکاربرد از جمله توزیع بتا، دوجمله‌ای کوشی، کای‌دو، نمایی،  $F$ ، گاما، هندسی، فوق هندسی، لوژستیک، لگ‌نرمال، دوجمله‌ای منفی، نرمال، پواسون،  $t$ ی استیودنت، یکنواخت و وایبول در R امکان‌پذیر است. هریک از این توزیع‌ها با یک کُد مخصوص به خود شناخته می‌شوند. و با افزودن حروفی به ابتدای کُد هر توزیع می‌توانیم توابع چگالی احتمال و جرم احتمال و توزیع احتمال و چندکی و تولید اعداد تصادفی برای آن توزیع را داشته باشیم.

علاوه بر این موارد هر توزیع پارامترهای مختص خود را دارد که بطور پیش‌فرض نیز یک مقدار از پیش تعیین‌شده نیز دارند و می‌بایست آن‌را نیز در نظر داشته باشیم.

ردیف	پارامترها در R	کُد توزیع	نماد ریاضی	نام توزیع
۱	shape1 = $\alpha$ , shape2 = $\beta$	beta	$B(\alpha, \beta)$	بتا
۲	size = $n$ , prob = $p$	binom	$Bin(n, p)$	دوجمله‌ای
۳	location = $\mu$ , scale = $\sigma$	cauchy	$C(\mu, \sigma)$	کوشی
۴	df = $\nu$ , ncp = $\lambda$	chisq	$\chi^2_{\mu, \lambda}$	کای دو
۵	rate = $\lambda$	exp	$E(\lambda)$	نمایی
۶	df1 = $\nu_1$ , df2 = $\nu_2$ , ncp = $\lambda$	f	$F_{\nu_1 \nu_2, \lambda}$	F
۷	shape = $\alpha$ , rate = $\frac{1}{\beta}$ , scale = $\beta$	gamma	$\Gamma(\alpha, \beta)$	گاما

ردیف	پارامترها در R	کُد توزیع	نماد ریاضی	نام توزیع
۸	$\text{prob} = p$	geom	$G(p)$	هندسی
۹	$m = M$ , $n = N - M$ , $k = n$	hyptr	$H(N, M, n)$	فوق هندسی
۱۰	$\text{location} = \mu$ , $\text{scale} = \lambda$	logis	$L(\mu, \lambda)$	لوژیستیک
۱۱	$\text{meanlog} = \mu$ , $\text{sdlog} = \sigma$	lnorm	$LN(\mu, \sigma^2)$	لگ نرمال
۱۲	$\text{size} = r$ , $\text{prob} = p$ , $\mu = \frac{r(1-p)}{p}$	nbinom	$NBin(r, p)$	دوجمله‌ای منفی

ردیف	پارامترها در R	کُد توزیع	نماد ریاضی	نام توزیع
۱۳	$\text{mean} = \mu, \text{sd} = \sigma$	norm	$N(\mu, \sigma^2)$	نرمال
۱۴	$\text{lambda} = \lambda$	pois	$P(\lambda)$	پواسون
۱۵	$\text{df} = \nu, \text{ncp} = \lambda$	t	$t_{\nu, \lambda}$	Tی استیودنت
۱۶	$\text{min} = a, \text{max} = b$	unif	$U(a, b)$	یکنواخت
17	$\text{shape} = \alpha, \text{scale} = \lambda$	weibull	$W(\alpha, \lambda)$	وایبول