

# تولید داده

(تولید داده در تولید داده در R)

---

تدوین: مرکز آماری خوارزمی

\*\*\*

[www.kharazmi-statistics.ir](http://www.kharazmi-statistics.ir)

مرکز آماری خوارزمی

---

## تولید داده در نرم افزار R

داده ها به دو صورت قابل تولید هستند، که به شرح زیر می باشند.

### ✓ ایجاد دنباله های منظم از داده ها

می خواهیم دنباله منظم از اعداد صحیح مثلا اعداد ۱ تا ۳۰ را در نرم افزار R تولید کنیم. برای این منظور باید یکی از عبارت های زیر را در نرم افزار بنویسید. نتیجه دستور یک آرایه با ۳۰ عنصر است. عملگر (( : )) در بین عملگر ها اولویت دارد.

$>x<-1:30$  (1)

$>s=1:30$  (2)

```
> x<-1:30
> x
[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
> s=1:30
> s
[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
```

خروجی نرم افزار به صورت تصویر بالا خواهد بود.

دستورات زیر را بنویسید و به تفاوت هایی که در جواب ها نمایش داده شده است دقت کنید.

$>1:10-2$

$>1:(10-2)$

```
> 1:10-2
[1] -1 0 1 2 3 4 5 6 7 8
> 1:(10-2)
[1] 1 2 3 4 5 6 7 8
```

در این عبارت ابتدا سری ۱ تا ۱۰ تولید شده سپس عدد ۲ از همه ی اعداد تولید شده کسر می گردد.

در این عبارت ابتدا عدد ۲ از ۱۰ کسر می شود سپس سری از ۱ تا ۸ تولید می شود.

همانطور که مشاهده می شود اعداد تولید شده در سری با توجه به تقدم در علائم متفاوت است.

## Seq()

تابع seq() می تواند دنباله ای از اعداد حقیقی را تولید کند.

```
>seq(2,8,0.7)
```

```
> seq(2, 8, 0.8)
[1] 2.0 2.8 3.6 4.4 5.2 6.0 6.8 7.6
```

اعداد تولید شده در نرم افزار

در این تابع، اولین عدد شروع دنباله، دومین عدد خاتمه دنباله و سومین عدد میزان افزایش در اعداد دنباله و فاصله دو عدد متوالی در دنباله را نشان می دهد.

همچنین می توان نوشت:

```
>seq(from=2,to=9,length=12)
```

```
> seq(from=2,to=9,length=12)
[1] 2.000000 2.636364 3.272727 3.909091 4.545455 5.181818 5.818182 6.454545 7.090909
[10] 7.727273 8.363636 9.000000
```

در این فرمان عدد اول عدد شروع دنباله دومین عدد، عدد انتهایی سری و عدد سوم تعداد اعدادی است که می خواهیم بین عدد اول و دوم تولید شود. (در این سری بین ۲ تا ۹، ۱۲ عدد تولید شده است.)

## c()

با تابع c() می توان اعداد مورد نظر را کنار هم قرار داد و به صورت یک دنباله در آورد. به c() تابع combine نیز می گویند. در این تابع اعدادی که مد نظر کاربر است قرار داده می شود این اعداد می تواند هیچ دنباله ی منظمی نباشد. و اعداد بستگی به آن عددهایی بستگی دارد که کاربر در نرم افزار می نویسد.

```
>c(1,3,4,6,3,75,34)
```

```
R Console
> c(1,3,4,6,3,75,34)
[1] 1 3 4 6 3 75 34
```

خروجی نرم افزار

خروجی نرم افزار دقیقاً اعدادی است کاربر وارد کرده است. این اعداد می تواند به صورت سری دنباله دار نوشته شود و یا نه.

## rep()

تابع دیگری تحت عنوان rep() وجود دارد که اولین آرگومان آن تابع تعریف شده برای اعداد نوشته شده و آرگومان دوم تعداد تکرار عناصر آرگومان اول است. عبارت زیر را در نرم افزار بنویسید و جواب نمایش داده شده در با تصویر زیر مقایسه کنید.

```
>rep(1:6,3)
```

تابع تعریف شده

تعداد تکرار

```
> rep(1:6,3)
[1] 1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6
```

خروجی نرم افزار

همانطور که مشاهده می شود دنباله سری ۱ تا ۶ سه بار تکرار می شود. آرگومان دوم نیز می تواند بردار باشد که تعداد عناصر آن برابر با تعداد عناصر بردار آرگومان اول است. در این حالت هر عنصر بردار آرگومان دوم، تعداد تکرار عناصر آرگومان اول را معین می کند. عبارت ترکیبی که ترکیبی از دو تابع در نرم افزار است را در نرم افزار بنویسید.

```
>rep(1:4,c(4,2,3,1))
```

میزان تکرار هر یک از اعداد در سری

```
> rep(1:4,c(4,2,3,1))
[1] 1 1 1 1 2 2 3 3 3 4
```

در عبارت بالا طبق تابع c() عدد اول ۴ بار عدد دوم ۲ بار عدد سوم ۳ بار و عدد چهارم ۱ بار تکرار می شود. عبارت زیر را در نرم افزار بنویسید.

```
>rep('statistics',3)
```

```
> rep('statistics',3)
[1] "statistics" "statistics" "statistics"
```

عبارتی که درون (( ' ' )) نوشته شده دقیقا به تعداد تکراری که کاربر در نرم افزار نوشته شده است در خروجی نهایی نمایش داده می شود.

✓ ایجاد دنباله های تصادفی از داده ها

**rfunc(n,p1,p2,..)**

این نوع دنباله توسط توزیع های آماری قابل تولید است. شکل عمومی این تابع به صورت  $rfunc(n,p1,p2,...)$  است که حرف  $r$  از واژه  $random$  به معنای تصادفی اخذ شده است. در آن ها  $n$  تعداد اعدادی است که قرار است تولید شود. حروف  $p1,p2, \dots$  مقادیر پارامترهای تابع را نشان می دهد. یکی از توابع تعریف شده تابع نرمال است که با  $norm$  نوشته شده است. عبارت زیر را در نرم افزار بنویسید.

```
>rnorm(3)
```

```
> rnorm(3)
[1] -1.18263633 -0.05095184  0.67640268
```

در اینجا ۳ عدد تصادفی با توزیع نرمال استاندارد (میانگین صفر و انحراف معیار یک) تولید شده است.

در این تابع می توان از عبارت های محاسباتی و توابع دیگر هم استفاده کرد و اعداد را تولید کرد. مثلا در عبارت پایین سه عدد با توزیع نرمال استاندارد تولید می شود و عبارت با عدد ۲ جمع می گردد.

```
> 2+rnorm(3)
[1] 0.6473731 2.4060358 2.7709341
```

در هر مرحله برای استفاده از گزینه ی  $help$  از عبارت  $??$  استفاده کنید. مانند زیر

```
>??data
```

در عبارت بالا در قسمت  $help$  تمام عبارت هایی که  $data$  دارد را در یک صفحه ی اینترنتی می دهد.

منبع:

۱. آشنایی با زبان محاسبات  $R$ ، سید سعید موسوی ندوشنی، پاییز ۱۳۹۱

۲.  $R$  for Beginners, Emmanuel Paradis, Institut des Sciences de l' \_ Evolution, 12<sup>th</sup> Septetber 2005