

Valores NA

Téllez Gerardo Rubén

27/3/2021

Valores NA

NA 'Not available' indican que las entradas correspondientes a las entradas a esos índices no existe.

Se van a usar cuando no conozcamos un valor.

Podemos obtener valores booleanos a partir de condicionales.

```
length(1:100) > 400
```

```
## [1] FALSE
```

```
length(1:100) > 4
```

```
## [1] TRUE
```

Para hacer que una posición cambie de valor en un índice, incluyendo un valor no asignado para el índice

```
x <- seq(3, 50, by=3.5)  
x[3]
```

```
## [1] 10
```

```
x[3] = 20  
x
```

```
## [1] 3.0 6.5 20.0 13.5 17.0 20.5 24.0 27.5 31.0 34.5 38.0 41.5 45.0 48.5
```

```
x[16] #NA
```

```
## [1] NA
```

```
x[16]=55  
x
```

```
## [1] 3.0 6.5 20.0 13.5 17.0 20.5 24.0 27.5 31.0 34.5 38.0 41.5 45.0 48.5 NA  
## [16] 55.0
```

Para hacer una suma a elementos de un vector

```
x
```

```
## [1] 3.0 6.5 20.0 13.5 17.0 20.5 24.0 27.5 31.0 34.5 38.0 41.5 45.0 48.5 NA
## [16] 55.0
```

```
x[2:5] = x[2:5] + 1
```

```
x
```

```
## [1] 3.0 7.5 21.0 14.5 18.0 20.5 24.0 27.5 31.0 34.5 38.0 41.5 45.0 48.5 NA
## [16] 55.0
```

Para agregar entradas a índices NA

```
x[length(x) + 5] = 100
```

```
x
```

```
## [1] 3.0 7.5 21.0 14.5 18.0 20.5 24.0 27.5 31.0 34.5 38.0 41.5
## [13] 45.0 48.5 NA 55.0 NA NA NA NA 100.0
```

Operar con un vector con NAs va probablemte a entregar un NA

```
cumsum(x)
```

```
## [1] 3.0 10.5 31.5 46.0 64.0 84.5 108.5 136.0 167.0 201.5 239.5 281.0
## [13] 326.0 374.5 NA NA NA NA NA NA NA
```

```
sum(x)
```

```
## [1] NA
```

Eliminar NAs

Para operar con un vector con NAs podemos usar un el parámetro `na.rm = TRUE`

```
sum(x, na.rm = TRUE)
```

```
## [1] 529.5
```

```
mean(x, na.rm = TRUE)
```

```
## [1] 33.09375
```

Para remover las entradas de los vectores que sean NA

```
xi = x
```

```
# xi = xi[-which(xi == NA)] Esto no funciona, porque el NA no es un valor en sí mismo, no se puede comp
xi_A <- xi[-which(is.na(xi))]
```

```
#También habría podido pedir las entradas done no son NA
```

```
x[!is.na(x)]
```

```
## [1] 3.0 7.5 21.0 14.5 18.0 20.5 24.0 27.5 31.0 34.5 38.0 41.5
## [13] 45.0 48.5 55.0 100.0
```

En estadística es común cambiar los NA por un valor, comunmente la Media \bar{x}/μ

```
xi_u = xi
xi_u[which(is.na(xi))] = mean(xi_A)
xi_u
```

```
## [1] 3.00000 7.50000 21.00000 14.50000 18.00000 20.50000 24.00000
## [8] 27.50000 31.00000 34.50000 38.00000 41.50000 45.00000 48.50000
## [15] 33.09375 55.00000 33.09375 33.09375 33.09375 33.09375 33.09375 100.00000
```

Drásticamente se puede omitir las posiciones de NAs, **NO RECOMENDADO PARA ESTADÍSTICA DESCRIPTIVA** para funciones estadísticas.

```
x_om = na.omit(x)
x_om
```

```
## [1] 3.0 7.5 21.0 14.5 18.0 20.5 24.0 27.5 31.0 34.5 38.0 41.5
## [13] 45.0 48.5 55.0 100.0
## attr("na.action")
## [1] 15 17 18 19 20
## attr("class")
## [1] "omit"
```

```
cumsum(x_om)
```

```
## [1] 3.0 10.5 31.5 46.0 64.0 84.5 108.5 136.0 167.0 201.5 239.5 281.0
## [13] 326.0 374.5 429.5 529.5
```

```
sum(x_om)
```

```
## [1] 529.5
```

Para eliminar los attr que arrastran los na.omit u otros

```
attr(x_om, "na.action") = NULL
```

También es raro necesitar de eliminar los atributos.

Operaciones acumulativas

No puedo usar el parámetro `na.rm = TRUE` para funciones acumulativas como `cumsum`, en cambio debo otorgar el vector libre de NAs

```
cumsum(x[!is.na(x)])
```

```
## [1] 3.0 10.5 31.5 46.0 64.0 84.5 108.5 136.0 167.0 201.5 239.5 281.0  
## [13] 326.0 374.5 429.5 529.5
```