

# Examples of Random Draws in R

Fan Wang

2020-04-01

## Drawing Random Numbers

Go back to [fan's REconTools](#) Package, [R4Econ](#) Repository, or [Intro Stats with R](#) Repository.

```
options(knitr.duplicate.label = 'allow')

library(tidyverse)
library(tidyr)
library(knitr)
library(kableExtra)
# file name
st_file_name = 'fs_rand_draws'
# Generate R File
try(purl(paste0(st_file_name, ".Rmd"), output=paste0(st_file_name, ".R"), documentation = 2))
# Generate PDF and HTML
# rmarkdown::render("C:/Users/fan/R4Econ/support/rand/fs_rand_draws.Rmd", "pdf_document")
# rmarkdown::render("C:/Users/fan/R4Econ/support/rand/fs_rand_draws.Rmd", "html_document")
```

## Discrete Random Draws

### Draw Random Subset of Sample

- r random discrete

We have a sample of  $N$  individuals in some dataframe. Draw without replacement a subset  $M < N$  of rows.

```
# parameters, it_M < it_N
it_N <- 10
it_M <- 5

# Draw it_m from indexed list of it_N
set.seed(123)
ar_it_rand_idx <- sample(it_N, it_M, replace=FALSE)

# dataframe
df_full <- as_tibble(matrix(rnorm(4,mean=0,sd=1), nrow=it_N, ncol=4)) %>% rowid_to_column(var = "ID")

# random Subset
df_rand_sub_a <- df_full[ar_it_rand_idx,]

# Random subset also
df_rand_sub_b <- df_full[sample(dim(df_full)[1], it_M, replace=FALSE),]

# Print
# Display
```

```
kable(df_full) %>%  
  kable_styling_fc_wide()
```

ID
V1
V2
V3
V4
1
0.1292877
0.4609162
0.1292877
0.4609162
2
1.7150650
-1.2650612
1.7150650
-1.2650612
3
0.4609162
0.1292877
0.4609162
0.1292877
4
-1.2650612
1.7150650
-1.2650612
1.7150650
5
0.1292877
0.4609162
0.1292877
0.4609162
6
1.7150650
-1.2650612
1.7150650

-1.2650612

7

0.4609162

0.1292877

0.4609162

0.1292877

8

-1.2650612

1.7150650

-1.2650612

1.7150650

9

0.1292877

0.4609162

0.1292877

0.4609162

10

1.7150650

-1.2650612

1.7150650

-1.2650612

```
kable(df_rand_sub_a) %>%  
  kable_styling_fc_wide()
```

ID

V1

V2

V3

V4

3

0.4609162

0.1292877

0.4609162

0.1292877

10

1.7150650

-1.2650612

1.7150650

-1.2650612

2

1.7150650

-1.2650612

1.7150650

-1.2650612

8

-1.2650612

1.7150650

-1.2650612

1.7150650

6

1.7150650

-1.2650612

1.7150650

-1.2650612

```
kable(df_rand_sub_b) %>%  
  kable_styling_fc_wide()
```

ID

V1

V2

V3

V4

5

0.1292877

0.4609162

0.1292877

0.4609162

3

0.4609162

0.1292877

0.4609162

0.1292877

9

0.1292877

0.4609162  
 0.1292877  
 0.4609162  
 1  
 0.1292877  
 0.4609162  
 0.1292877  
 0.4609162  
 4  
 -1.2650612  
 1.7150650  
 -1.2650612  
 1.7150650

**Random Subset of Panel** There are  $N$  individuals, each could be observed  $M$  times, but then select a subset of rows only, so each person is randomly observed only a subset of times. Specifically, there are 3 unique students with student ids, and the second variable shows the random dates in which the student showed up in class, out of the 10 classes available.

```
# Define
it_N <- 3
it_M <- 10
svr_id <- 'student_id'

# dataframe
set.seed(123)
df_panel_rand <- as_tibble(matrix(it_M, nrow=it_N, ncol=1)) %>%
  rowid_to_column(var = svr_id) %>%
  uncount(V1) %>%
  group_by(!sym(svr_id)) %>% mutate(date = row_number()) %>%
  ungroup() %>% mutate(in_class = case_when(rnorm(n(), mean=0, sd=1) < 0 ~ 1, TRUE ~ 0)) %>%
  filter(in_class == 1) %>% select(!sym(svr_id), date) %>%
  rename(date_in_class = date)

# Print
kable(df_panel_rand) %>%
  kable_styling_fc_wide()
```

student\_id  
 date\_in\_class  
 1  
 1  
 1  
 2  
 1

8  
1  
9  
1  
10  
2  
5  
2  
8  
2  
10  
3  
1  
3  
2  
3  
3  
3  
3  
4  
3  
5  
3  
6  
3  
9