

LAB 2

Random number Generation

How does R get Random numbers to sample?

Truly random numbers are expensive

So R uses a “pseudo”random number generator.

This kind of generation is a deterministic (depends on the starting value) process that is almost the same as a true random process

The **seed** is the starting value that determines this sequence.

Why do we need to set a seed?

We want some way to “save” our results of a random process.

We want to use **the `set.seed()` function to set a seed.**

The **`set.seed()`** function sets the starting number used to generate a sequence of random numbers.

This means that you will get the same result if you run the same process again (if you keep the seed set at the same number).

Example:

```
```{r}  
set.seed(06041997)
runif(1)
```
```

```
[1] 0.4464028
```

we set the seed
to be any
number (I
chose my
birthday)

We want to keep this
result because if we
run it without (or with
a different seed) the
seed it might change.

Random draw from
a uniform distribution (min=0, max=1)

load()

load("x") function allows r to read a compressed file/url

```
load(url("https://stat.duke.edu/~mc301/data/ames.RData"))
```



For this lab we want to load a file from a url,
x. url("x") then we load.

```
load(url("x"))
```

Histogram

```
hist(x, breaks= N)
```

`x` <- is a vector of numeric values that we want the histogram to be built around

`N` <- “a single number giving the number of cells for the histogram”

```
load(url("https://stat.duke.edu/~mc301/data/ames.RData"))  
area <- ames$area  
hist(area, breaks = 100)
```

summary()

summary(x) gives us a summary of a desired vector x

Gives the output :

Min., 1st Qu., Median, Mean, 3rd Qu, Max,

sample()

x <- vector we would like to sample from

N <- how many values do we want in the sample

```
sample(x, size=N,replace = TRUE,  
        prob = c(0.2, 0.8))
```

The output is a new vector of size N