

# Stat 312: Lecture 01

## R Basics

Moo K. Chung  
mchung@stat.wisc.edu

September 2, 2004

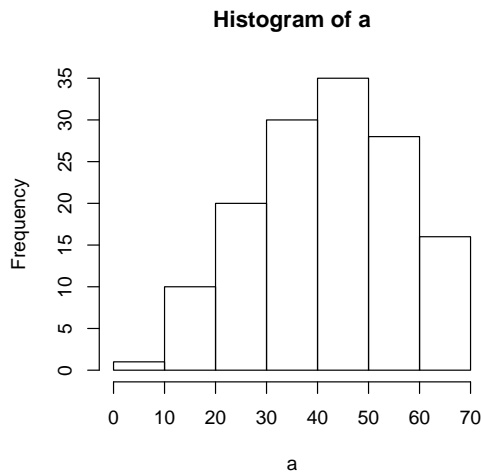


Figure 1: Histogram for binge drinker data

1. *Data loading.* Let's illustrate basic R commands by following example 1.5 about binge drinking in college. Data is the percentage of binge drinkers on 140 campuses across the United States.

```
> library(Devore6)
> data(xmp01.05)
> attach(xmp01.05)
> xmp01.05
  bingePct
1         4
2        11
3        13
.
.
.
.
138       67
139       67
140       68
```

2. *Histogram.* Let's see some basic R commands.

```
> a<-bingePct
> mean(a)
[1] 42.33571
> var(a)
[1] 205.8361
> sd(a)
[1] 14.34699
> hist(a)
```

To find out more about `hist` command, use `> help(hist)`. It will display a new window with detailed information about `hist`.

3. Binge drinking percentage can be modeled statistically. Let  $X_i$  is the binge drinking percentage at  $i$ -th campus which is distributed normally with mean 40 and standard deviation of 14, i.e.  $X_i \sim N(40, 14^2)$ . Since binge drinking percentages in different campuses should be independent,  $X_i$ 's should be assumed to be independent random variables.
4.  $X_1, \dots, X_n$  form a *random sample* if  $X_i$  are independent and identically distributed random variables. A *statistic* is a random variable. Hence *sample mean*  $\bar{X} = \sum_{i=1}^n X_i/n$  is a statistic. It will be distributed as

$$\bar{X} \sim N(40/140, 14^2/140).$$

*Note.* Read section 5.4, 5.5 and do Ex.5.48. 5.52. 5.63. Lecture 2-4 will be based on sections 6.1-6.2.