Introduction to Engineering Statistics and R

Engineering Statistics (IMSE 4410) Spring 2015. Copyright 2013-2015 by Timothy Middelkoop License CC by SA 3.0

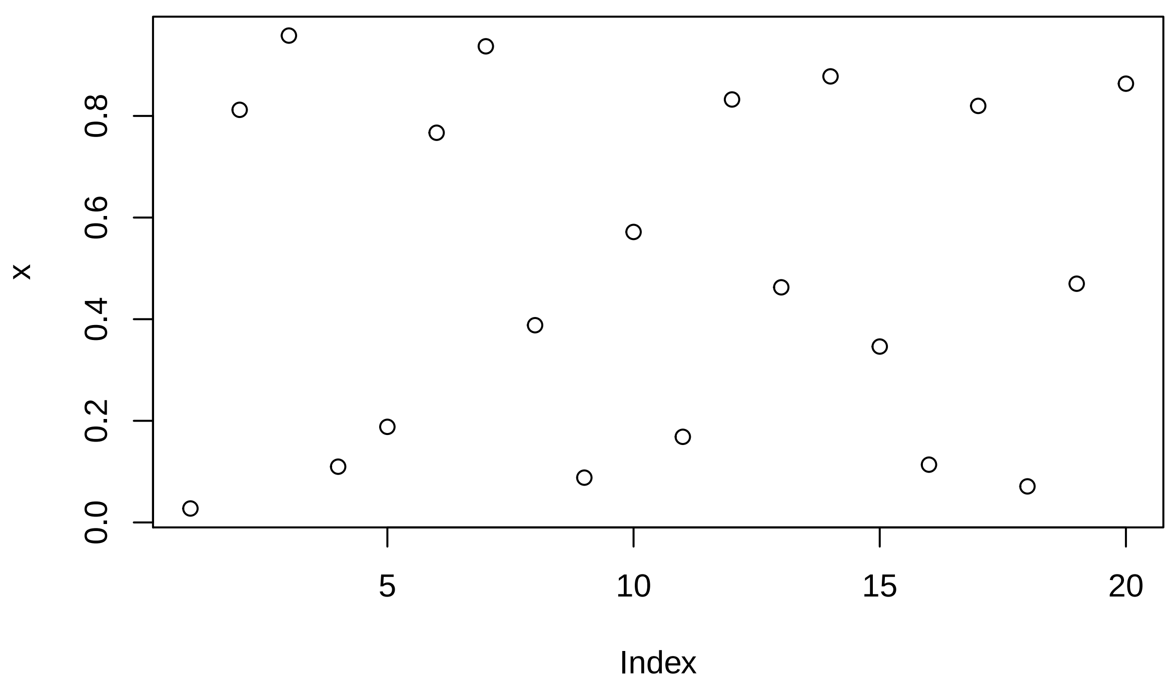
Review of Continuous Random Variables (Chapter 4, Week 3)

Uniform Random Variables.

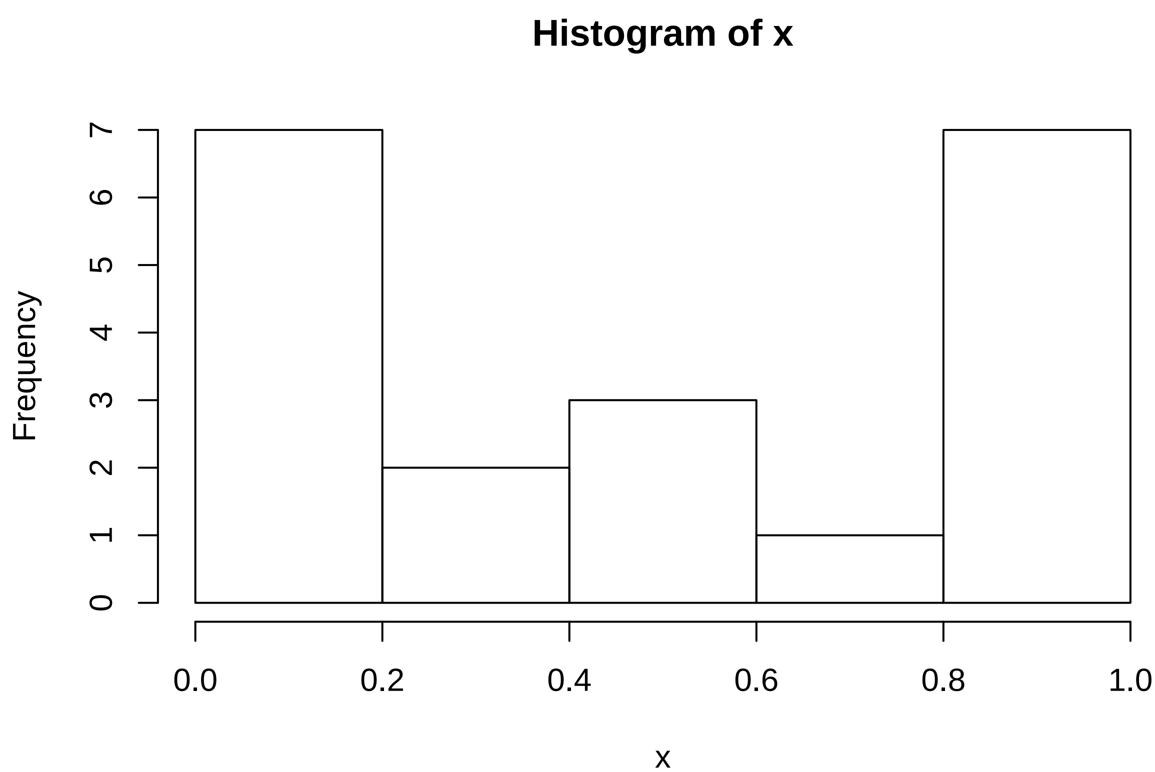
```
# What is a "random" number.  Dice is discrete, uniform. Here is the continuous version.
# 100 (r)andom (unif)orm numbers stored in x
x <- runif(20) ; x
```

```
##  [1] 0.02744114 0.81204100 0.95811624 0.10973709 0.18810256 0.76692357
##  [7] 0.93703153 0.38813517 0.08813907 0.57167185 0.16847882 0.83240147
## [13] 0.46269525 0.87783930 0.34613222 0.11372462 0.81972490 0.07092825
## [19] 0.46984328 0.86353210
```

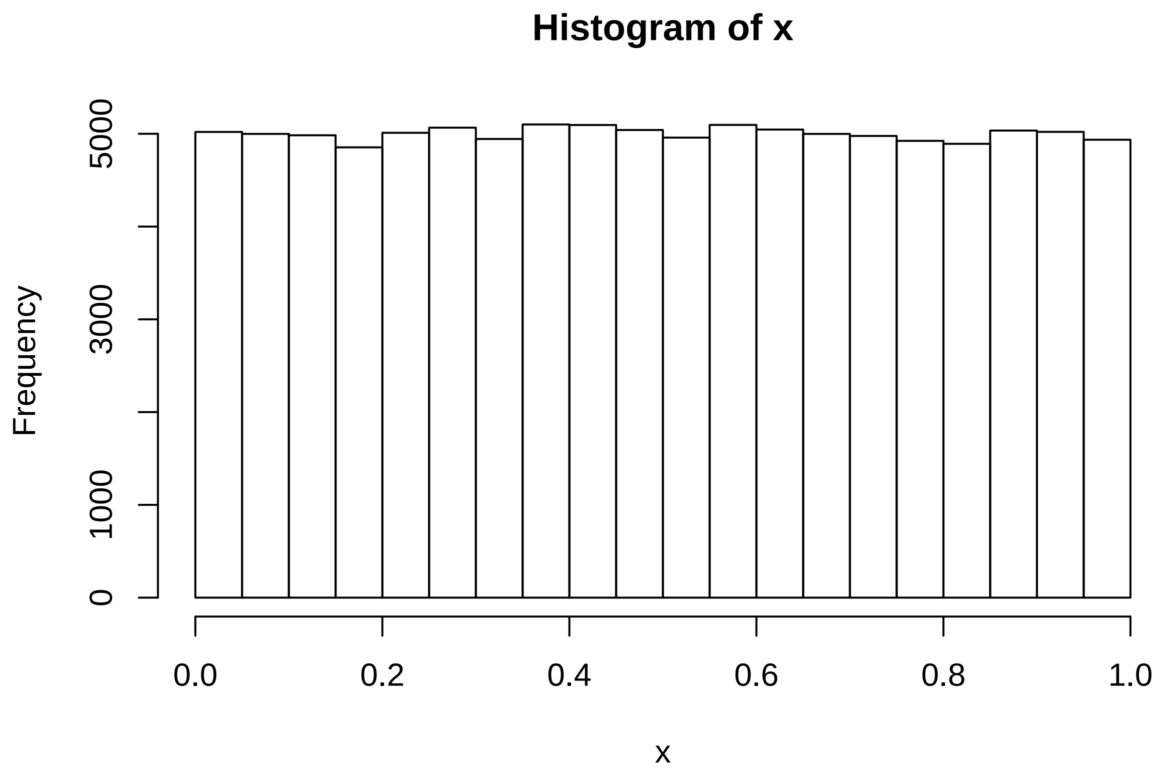
```
# visualize
plot(x)
```



```r
hist(x)
```



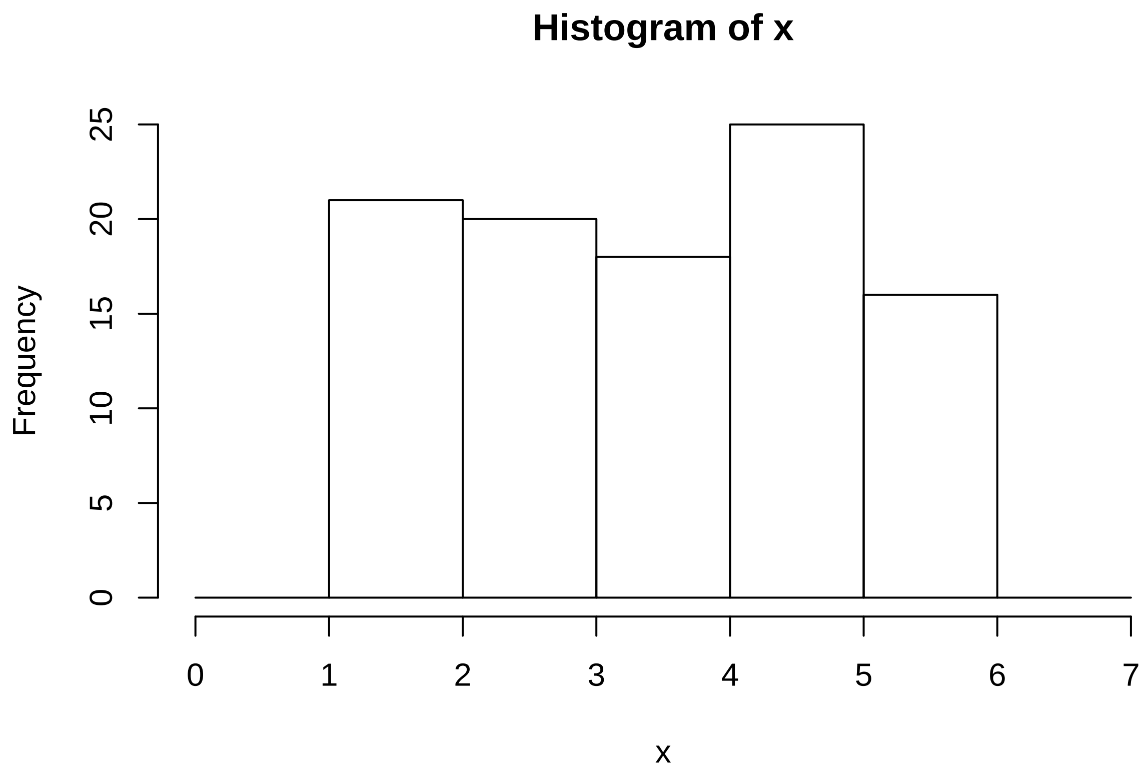

```r
# Close, lets make the numbers a bit larger
x <- runif(100000)
h <- hist(x)
```




```r
summary(x)
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## 0.0000387 0.2515000 0.4987000 0.4997000 0.7483000 1.0000000
```

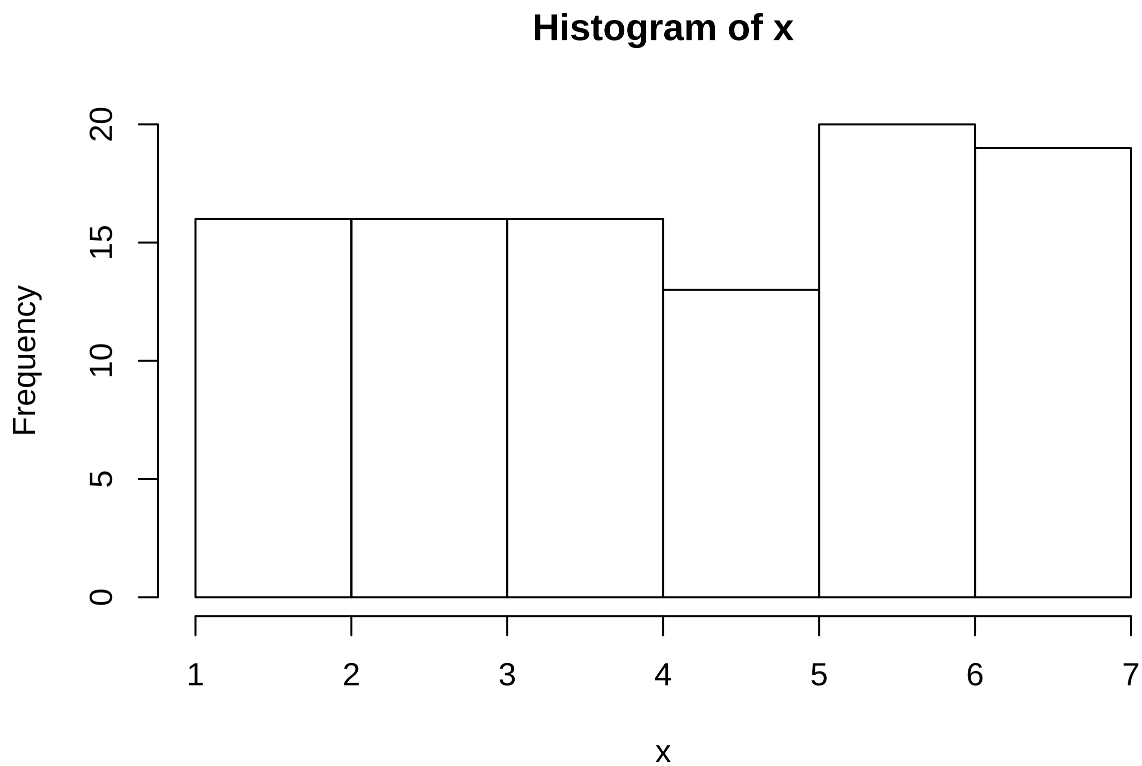
```r
# Looks good

# This histogram is the discrete probably density function.
# Lets roll a continuous 6 sided dice
x <- runif(100,1,6)
hist(x,breaks=0:7)
```




```r
# breaks are done explicitly
0:7
```

```
## [1] 0 1 2 3 4 5 6 7
```

```r
# Is is that what you expected?
x <- runif(100,1,7)
hist(x)
```




```r
# probability for discrete version is
1/6
```

```
## [1] 0.1666667
```

```r
# (d)ensity for the (unif)orm version (d=P(X==x) [density])
dunif(1:6,1,7)
```

```
## [1] 0.1666667 0.1666667 0.1666667 0.1666667 0.1666667 0.1666667
```

```r
# verify the property that the area under the curve=1
(1/6)*(7-1)
```

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

```r
# ok so we did not need a calculator but you get the point.
# What is the probability of rolling a 1 or less?
# What is the probability of rolling a 6 or less?
punif(1,1,7)
```

```
## [1] 0
```

```r
punif(7,1,7)
```

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

```r
# How about between 2 and 3

# Sound familiar?  What about the reverse, at what is P[x<=q]=0.50?
qunif(0.50,1,7)
```

```
## [1] 4
```

```r
# Look at the chart, does this make sense?
```

Review of R functions:

- runif(n,...) - random values drawn from the distribution.
- dunif(x,...) - the fx (probably density function - pdf) values
- punif(q,...) - the Fx (cumulative distribution function - cdf) values P(X<=x)
- qunif(p,...) - the inverse cdf values (quantile), find q given P(X<=q)==p

Random Normal Distribution

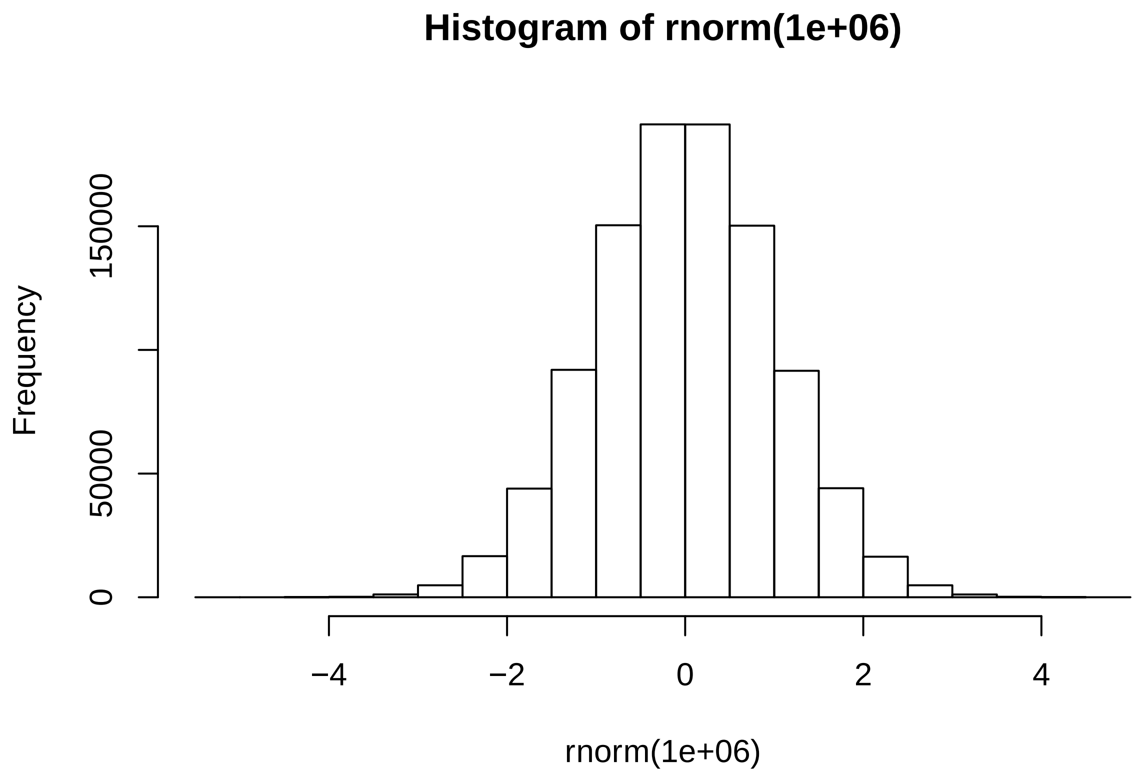
$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty$$

parameters of μ and σ or $N(\mu, \sigma)$

$$E(X) = \mu, V(X) = \sigma^2$$

- Standard norm, so we can use the tables, is $\mu = 0$, $\sigma^2 = 1$
- Random variable for the standard norm is Z, cumulative function F(z)=P(Z<=z)

```r
# The normal distribution is the rnorm series
hist(rnorm(1e6))
```




```r
# What is P(Z<=1.5)
# F(1.5)
# Tables... 0.933139
# or use
pnorm(1.5)
```

```
## [1] 0.9331928
```

```r
# Z=(X-mu)/sigma
mu <- 5 ; sigma <- 3
# find P(X<=5.5) should be close to zero and positive.
pnorm((5.5-mu)/sigma)
```

```
## [1] 0.5661838
```

```r
# [1] 0.5661838
# Note the use of the solution in the code.  This is called a "Known Good Solution"
# No need to convert to standard norm
pnorm(5.5,mean=mu,sd=sigma)
```

```
## [1] 0.5661838
```

Exponential Distribution

The other real useful distribution is the exponential distribution.

- $\mu = 1/\lambda$, $var = 1/\lambda^2$
- Memoryless
- Units must be the same.
- $1/\lambda$ is the mean interarrival time (time between events)
- λ is the mean arrival rate (events/time)

$$f(x) = \lambda e^{-\lambda x}, 0 \leq x < \infty$$

$\mu = E(X) = \frac{1}{\lambda}$ and $\sigma^2 = V(X) = \frac{1}{\lambda^2}$

It is important to use consistant units for X and λ.

```
# Example: 5 students per hour (1 every 12 min, 1 every .2 hour)
x <- rexp(1000,5)
hist(x)
# Look at the results
mean(x)
```

```
## [1] 0.2037005
```

```
var(x)
```

```
## [1] 0.04178562
```

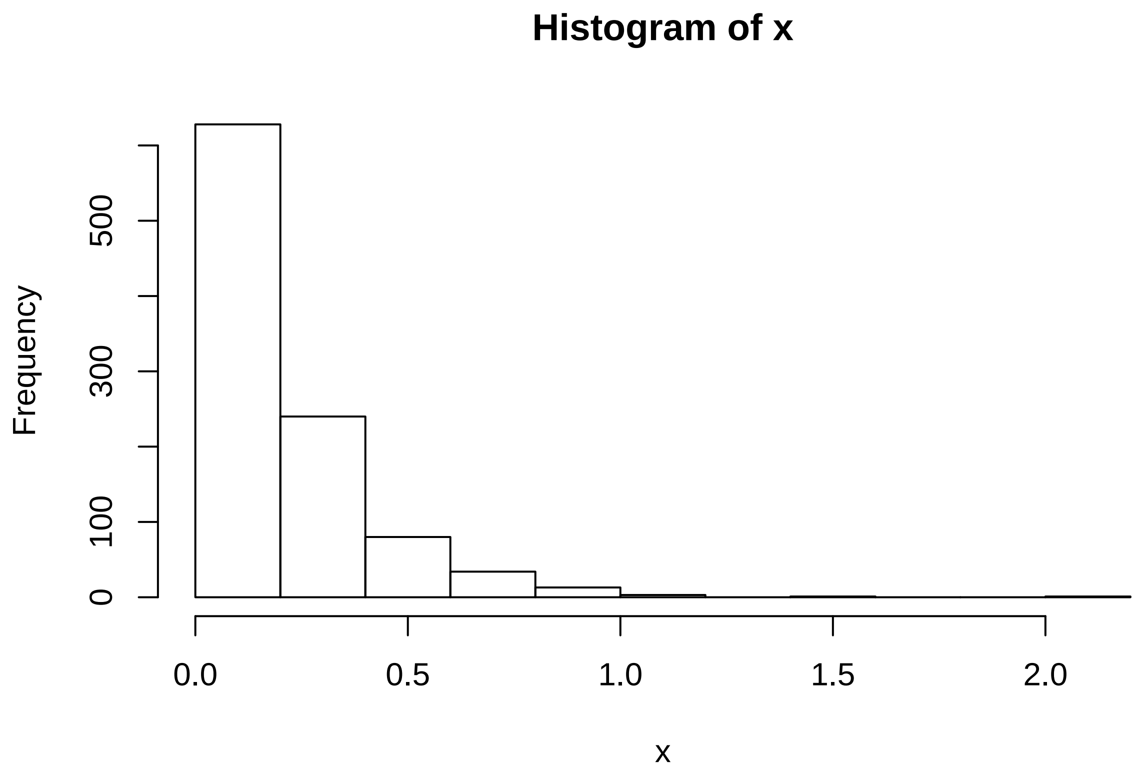
```
# Now from the tables.
# What is the probablility that a student will come in in the next 6 min
# This is .1 hours.
# P(X<=0.1)
pexp(.1,5)
```

```
## [1] 0.3934693
```

```
pexp(.2,5)
```

```
## [1] 0.6321206
```

```
## Does this make sence?
n <- length(x)
h <- hist(x)
```




```r
names(h)
```

```
## [1] "breaks"   "counts"   "density"  "mids"     "xname"    "equidist"
```

```r
cumsum(h$counts/n)
```

```
##  [1] 0.628 0.868 0.948 0.982 0.995 0.998 0.998 0.999 0.999 0.999 1.000
```