RESEARCH SCHOOL OF FINANCE, ACTUARIAL STUDIES AND APPLIED STATISTICS

Solutions to R Worksheet 3.

Q1(a)

```
> chivals <- rchisq(1000,df=18)
```

Q1(b)

```
> hist(chivals)
```

Q1(c)

```
> min(chivals)
[1] 4.895851

> max(chivals)
[1] 44.4928
```

Q1(d)

```
> hist(chivals,breaks=seq(floor(min(chivals)),floor(max(chivals)+1),1))
> hist(chivals,breaks=seq(floor(min(chivals)),floor(max(chivals)+1),0.5))
> hist(chivals,breaks=seq(floor(min(chivals)),floor(max(chivals)+1),0.25))
```

Q1(e)

```
> for(i in 1:24) {
+ chivals_chivals+rchisq(1000,df=18)
+ }
> chiavg <- chivals/25
> hist(chiavg, breaks=seq(floor(min(chiavg)),floor(max(chiavg)+1),0.5))
> mean(chiavg)
[1] 17.99582

#  The mean should be equal to 18, the degrees of freedom (which is the
#  expectation of a chi-squared random variable)

> sqrt(var(chiavg))
[1] 1.201205

# The variance of these values should be 36/25 = 1.44, twice the degrees of
# freedom (which is the variance of a chi-squared random variable) over the
# number of summands in the average.  Thus, the standard deviation out to
# be 1.2, which is the square root of 1.44.
```

Q1(f)

```
> grid <- seq(floor(min(chiavg)),floor(max(chiavg)+1),0.1)
> normals <- dnorm(grid,18,1.2)
> hist(chiavg, breaks=seq(floor(min(chiavg)),floor(max(chiavg)+1),0.5))
> lines(grid,normals)
> lines(grid,1000*0.5*normals)
```

Q1(g)

```
> qqnorm(chiavg)
```

Q2

```
# The function was given in the worksheet.
# To test it, we might look at the "worksheet2.women2" dataframe.

> worksheet2.women <- as.data.frame(worksheet2.women)
> attach(women.df)
> influence(height,1)
[1]  0.4533333 -4.2446111
> influence(height,2)
[1]  0.1422222 -6.3290556
> influence(height,3)
[1] -0.9355556  3.2901111
> influence(height,4)
[1]  1.731111 27.156778
> influence(height,5)
[1] -0.7911111  0.4842778
> influence(height,6)
[1]  0.008888889 -6.555722222
> influence(height,7)
[1]  0.1311111 -6.3632222
> influence(height,8)
[1] -1.002222  4.743444
> influence(height,9)
[1]  0.2866667 -5.6321111
> influence(height,10)
[1] -0.02444444 -6.54988889

# Looks like point 4 is one to look out for

> influence(weight,1)
[1] 1.303333 8.944000
> influence(weight,2)
[1] -0.1411111 -9.9421111
> influence(weight,3)
[1] -0.3855556 -8.4937778
> influence(weight,4)
[1]  0.5588889 -6.6521111
> influence(weight,5)
[1] -0.7188889 -4.3521111
> influence(weight,6)
[1] -0.7855556 -3.2237778
> influence(weight,7)
[1] -0.2966667 -9.1760000
> influence(weight,8)
[1] -1.141111  4.482889
> influence(weight,9)
[1] -0.430 -8.086
> influence(weight,10)
[1]  2.036667 36.499000
> influence(weight,c(1,10))
[1]  3.75750 60.55346
> influence(weight,c(1,2))
[1]  1.307500 -1.396536
> influence(weight,c(2,10))
[1]  2.13250 30.18918

# Looks like point 10 is the one to look out for here, and in combination
# with point 1 they seem to have a rather large effect.
```