Joseph Blubaugh
Stat 630
Hw07

1) 4.6.1

$$Let X_1 = N(3, 2^2), X_2 = N(-8, 5^2)$$

$$U = x_1 - 5x_2, V = -6x_1 + Cx_2$$

$$U = N(3 - 5(-8), 4 - 5^2 5^2) = N(43, 629)$$

$$\begin{aligned} V &= N(-6(3) + c(-8), -6^2(4) + C^2(25)) \\ &= N(-18 - 8C, 144 + 25C^2) \end{aligned}$$

2) 4.6.3

$$Let X = N(3, 5), Let Y = N(-7, 2)$$

$$\chi^2(C_5) = C_1(x + C_2)^2 + C_3(Y + C_4)^2$$

$$\begin{aligned} C_1(x + C_2)^2 &= C_1(\frac{3 + C_2}{\frac{\sqrt{1}}{C_1}}) = 0, C_2 = -3 \\ &= var(\sqrt{C_1(5 - 3)}) = 1, C_1 = 1/2 \end{aligned}$$

$$\begin{aligned} C_3(Y + C_4)^2 &= \frac{Y + C_4}{\frac{\sqrt{1}}{C_3}} = 0, C_4 = -7 \\ &= var(\sqrt{C_3(2 + 7)}) = 1, C_3 = 1/3 \end{aligned}$$

$$\frac{1}{2}(x - 3)^2 + \frac{1}{3}(y + 7)^2 = \chi^2(C_5), C_5 = 2$$

3) 4.6.5 expand chi-square from N(0,1)

$$x = \chi^2(n), y = \chi^2(m)$$

$$\begin{aligned} x &= z_1^2 + z_2^2 + ... + z_n^2, \chi^2(z = 1) \\ y &= t_1^2 + t_2^2 + ... + t_n^2, \chi^2(t = 1) \\ x + y &= \chi^2(z + t) \end{aligned}$$

4) 4.6.6

$$
\begin{aligned}
F(n,3n) &= C\frac{X_1^2+X_2^2+...+X_n^2}{X_{n+1}^2+X_{n+2}^2+...+X_{4n}} \\
C &= \frac{n}{3n} \\
&= \frac{1}{2n}
\end{aligned}
$$

5) 4.6.7 We need to divide the square root by n...

$$
\begin{aligned}
t(n) &= C\frac{X_1}{\sqrt{X_2^2+...+X_n^2+x_{n+1}^2}} \\
C &= \frac{1}{\sqrt{\frac{1}{n}}}
\end{aligned}
$$

6) 4.6.8

$$X = N(3,5), Y = N(-7,2)$$

$$t(C_6) = \frac{C_1(x+C_2)^{C_3}}{(Y+C_4)^{C_5}}$$

$$
\begin{aligned}
C_1(3+C_2)^{C_3} &= 0, C_2 = -3 \\
\sqrt{C_1(5-3)^{C_3}} &= 1, C_3 = 2 \\
\sqrt{C_1}2 &= 1, C_1 = 1/4
\end{aligned}
$$

$$
\begin{aligned}
(-7+C_4)^{C_5} &= 0, C_4 = 7 \\
\sqrt{(2+7)^{C_5}} &= 1, C_5 = 0
\end{aligned}
$$

$$t(C_6) = 1, C_6 = 1$$

7) 4.6.10

a) chi-square -> z ~ X^2(1)

b) chi-square -> z ~ X^2(5)

c) t-distribution -> z ~ t(3)

d) F-distribution -> z ~ F(3, 1)

e) F-distribution -> z ~ F(30, 70)

8) 4.6.12

a) t-distribution -> z ~ t(20)

b) t-distribution -> z ~ t(30)

c) F-distribution -> z ~ F(20,30)

d) F-distribution -> z ~ F(20,30)

e) chi-square -> X^2(20)

9) 5.1.11 ...

```
y.fun = function(x) x^4 + 2*x^3 - 3

# random sample of 10k
rand.samp = rnorm(10000)
y.results = y.fun(rand.samp)

length(which(y.results >=1 & y.results <= 2))/length(y.results)
```

```
## [1] 0.0185
```

10) 5.3.11

$$I = N(10, 2), II = N(8, 3)$$

$$\psi = ln(\frac{\theta}{(1-\theta)})$$
$$e^{\psi} = \frac{\theta}{1-\theta}$$

$$\frac{1-\theta}{\theta} = \frac{1}{e^{\psi}}$$
$$\frac{1}{\theta} - 1 = \frac{1}{e^{\psi}}$$
$$\frac{1}{\theta} = \frac{1}{e^{\psi}} + 1$$
$$\theta = e^{\psi} + 1$$

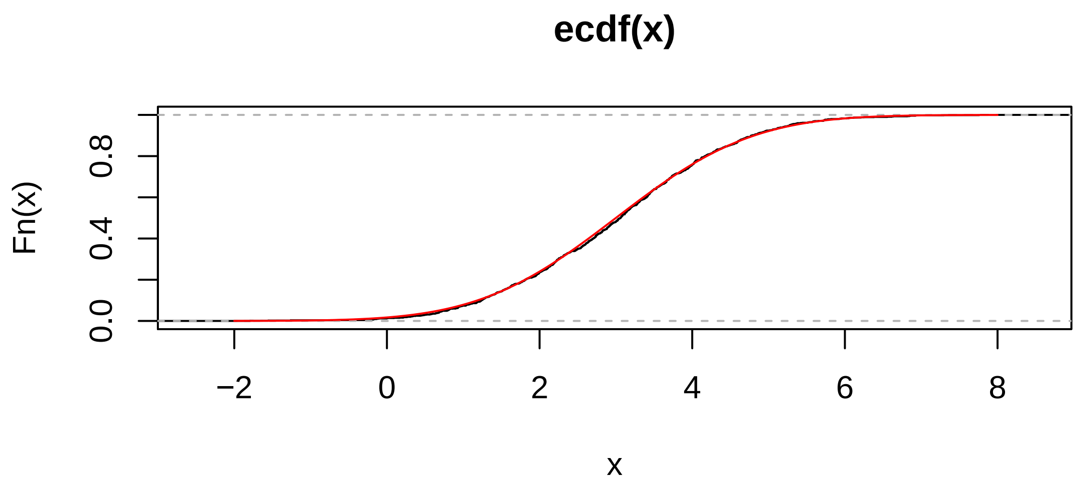
$$f_{\psi}(x_i) = (e^{(\psi+1)})^{x_i}(1 - e^{\psi+1})^{1-x_i}$$

11) 5.4.11 .

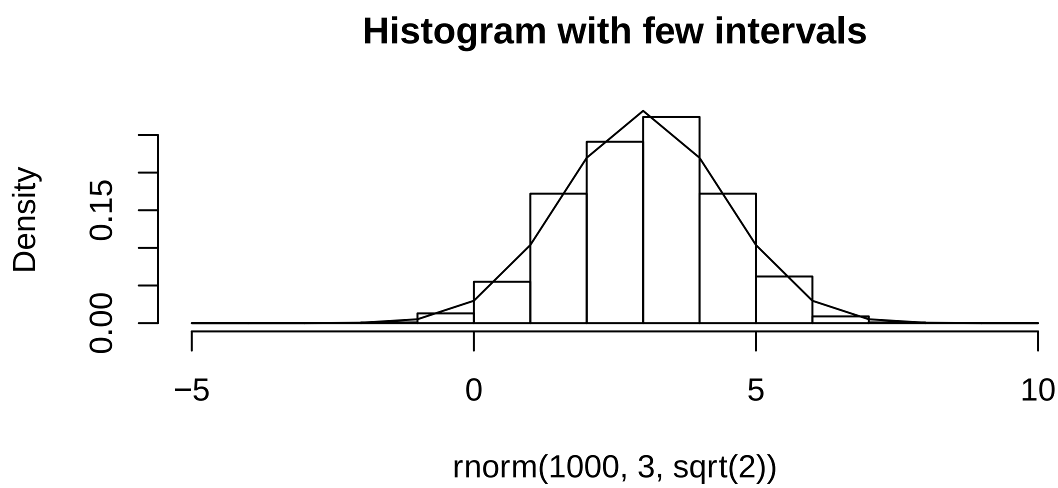
```
# a) random sample
sample = rnorm(1000, mean = 3, sd = sqrt(2))
mean(sample)
```

```
## [1] 3.024972
```

```r
plot.ecdf(sample)
x = seq(-2,8,length = 1001)
lines(x, pnorm(x, mean = 3,sd = sqrt(2)), col = 2)
```

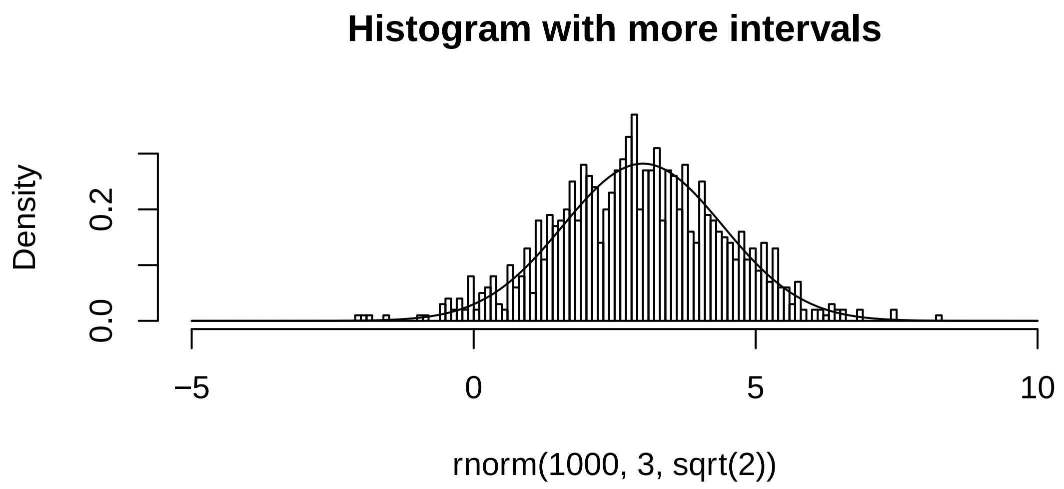



```r
# b)
x = seq(-5, 10, by = 1)
hist(x = rnorm(1000, 3, sqrt(2)), breaks = seq(-5, 10, by = 1), freq = FALSE,
     main = "Histogram with few intervals")
lines(x = x, y = dnorm(x = x, mean = 3, sd = sqrt(2)))
```




```r
# c)
x = seq(-5, 10, by = .1)
hist(x = rnorm(1000, 3, sqrt(2)), breaks = seq(-5, 10, by = .1), freq = FALSE,
```

```r
      main = "Histogram with more intervals")
lines(x = x, y = dnorm(x = x, mean = 3, sd = sqrt(2)))
```

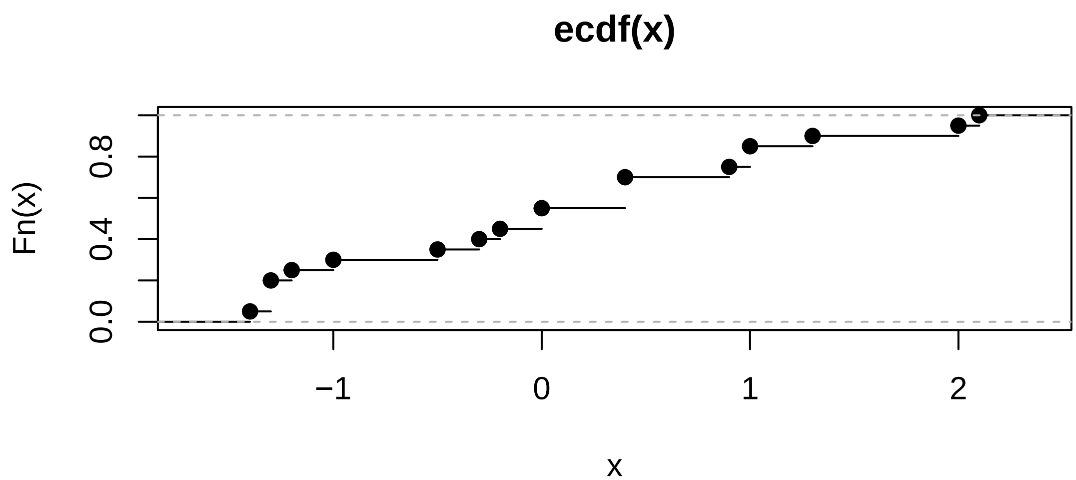


d) The first histogram is rigid and has relatively few intervals. This density curve is also no smooth because of the small number of intervals. The second histrogram is much smoother because of the increased number of intervals
e) The limit for the size of intervals is the number of observations in the dataset. With more observations you can afford to have smaller intervas and still see the distribution.

12) 5.5.5 …

```r
data = sort(c(1, -1.2, .4, 1.3, -.3, -1.4, .4, -.5, -.2, -1.3, 0, -1, -1.3, 2, 1,
              .9, .4, 2.1, 0, -1.3))

plot.ecdf(data)
```



```r
# median
(data[10] + data[11]) / 2
```

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

```r
# 1st quartile with type 5
(data[5] + data[6]) / 2
```

```
## [1] -1.1
```

```r
# 3rd quartile with type 5
(data[15] + data[16]) / 2
```

```
## [1] 0.95
```

```r
# interquartile range
((data[15] + data[16]) / 2) - ((data[5] + data[6]) / 2)
```

```
## [1] 2.05
```

```r
# F(1)
length(data[data < 1])/20
```

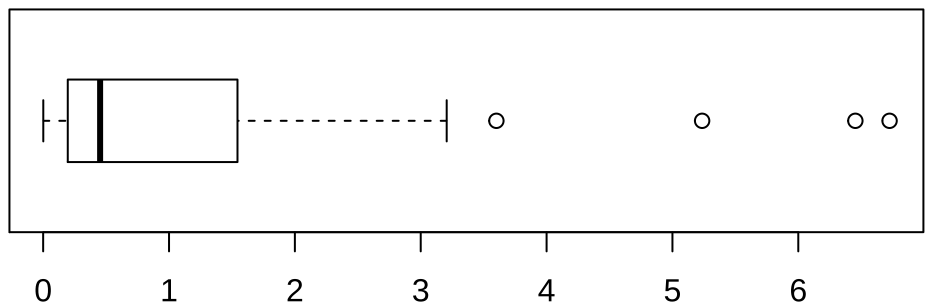
```
## [1] 0.75
```

13) 5.5.19 ...

```r
(data = rchisq(50, 1))
```

```
##  [1] 0.126070046 1.633093330 0.249120406 0.184087898 0.315549151
##  [6] 0.678520344 0.660400077 0.799511560 1.513001051 5.235729463
## [11] 0.769231934 0.027532259 3.206296561 0.744007784 3.600986463
## [16] 0.330978015 2.155509089 0.609662063 3.091678335 1.648773846
## [21] 2.611635546 0.072618807 0.313740391 1.123052376 0.451098599
## [26] 0.195615979 0.081296178 0.208874296 1.060828139 0.161120586
## [31] 0.261541563 0.023736789 0.332364327 0.571342446 0.123161382
## [36] 0.374331898 0.327544360 0.245169887 0.032699105 0.217284208
## [41] 0.087259510 2.944778871 0.454120433 6.725145703 6.453255435
## [46] 1.544064482 0.001066974 1.236687076 0.027784327 1.843738495
```

```r
# a)
boxplot(data, horizontal = T)
```



b) This plot is skewed heavily to the right
c) Max, 1.5*IQR, and quartiles would give a good indication of the distribution because it would show how the data are distributed far to the right of the median.

14) 5.5.20 ...

```
data = sort(rnorm(50, mean = 4, sd = sqrt(1)))

# a)
x = .9 * 50
(data[x] + data[x+1]) / 2
```

```
## [1] 5.218644
```

```
# b)
z = qnorm(p = .9, mean = 0, sd = sqrt(1))
mean(data) + z
```

```
## [1] 5.176847
```

c) I would prefer the estimate of the 90th percentile from part a because its an extreme number relative to the normal distribution and seems unlikely that if chosen at random, it would belong to that distribution.