STAT390: Homework 4

Ken

2023-10-27

Question 1

a). Geometric Distribution. It has one parameter, p, which is the probability of winning at least one of the lotteries. The value we pass in as p would be P("win one") = 1 - P("win neither") = (1 - (1 - $p1$)(1 - $p2$))

b). The expected number of times you would need to play for a win would be 1/(1 - (1 - $p1$)(1 - $p2$)).

Question 5

a).

```
pnorm(65, mean = 50, sd = 10) - pnorm(45, mean = 50, sd = 10)
```

```
## [1] 0.6246553
```

b).

```
top3 <- qnorm(.97, mean = 0, sd = 1)
top3 * 10 + 50
```

```
## [1] 68.80794
```

c).

```
mean <- qnorm((.9 - .05) / 2)
sd <- (qnorm(.9) - qnorm(.05)) / (2 * qnorm(0.5))
mean
```

```
## [1] -0.1891184
```

```
sd
```

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

Question 6

Homework 4

(2) a) $P(X=4) = \dfrac{\binom{7}{4}\binom{12-7}{6-4}}{\binom{12}{6}} = \dfrac{35 \cdot 10}{924}$

$M=7 \quad n=6$

$N=12 \quad X=4$

$= \boxed{\dfrac{350}{924} = .3788}$

$1-(P(X=5)+P(X=6))$

$\dfrac{\binom{7}{5}\binom{12-7}{6-5}}{\binom{12}{6}} + \dfrac{\binom{7}{6}\binom{12-7}{6-6}}{\binom{12}{6}}$

$1-\left(\dfrac{21 \cdot 5}{924} + \dfrac{7 \cdot 1}{924}\right) = \boxed{\dfrac{812}{924} = .8788}$

b) $E(X) = 6 \cdot \dfrac{7}{12} = \dfrac{21}{6} \qquad p = \dfrac{7}{12}$

$r =$ cost of repairs

$r = 75x + 30(6-x)$

$= 45x + 180$

$E[r] = E[45x+180] = 45E[x]+180$

$45 \cdot \dfrac{21}{6} + 180 = \boxed{337.5 = E[r]}$

Figure 1: Homework 4 Page 1

$$SD(r) = \sqrt{Var(45x+180)} = \sqrt{Var(45x)}$$

$$= \sqrt{45^2\, Var(x)} \qquad \frac{12-6}{12-1} \cdot \frac{21}{6}\left(1 - \frac{7}{12}\right)$$

$$= \sqrt{45^2\left(\frac{6}{11} \cdot \frac{21}{6} \cdot \frac{5}{12}\right)} = \boxed{40.135 = SD(r)}$$

3 a) $$\frac{3}{4}\int_0^x 2x - x^2\, dx \qquad \frac{3}{4}\left(x^2 - \frac{1}{3}x^3\right)$$

$$P(X \le x) = f_x(x) = \boxed{\frac{(3-x)x^2}{4} = cdf_x}$$

b) $$E[x] = \int_{-\infty}^{\infty} x f(x)\, dx$$

$$\frac{3}{4}\int_0^2 2x^2 - x^3\, dx \qquad \frac{3}{4}\left(\frac{2}{3}x^3 - \frac{1}{4}x^4\right)\Big|_0^2$$

$$\frac{3}{4}\left(\frac{16}{3} - \frac{16}{4}\right) = \frac{3}{4}\left(\frac{64}{12} - \frac{48}{12}\right) = \frac{3}{4}\left(\frac{16}{12}\right) = \boxed{1 = E[x]}$$

$$Var(x) = E[x^2] - E[x]^2 \qquad \frac{3}{4}\int_0^2 x^2(2x - x^2)\, dx$$

$$\frac{3}{4}\int_0^2 2x^3 - x^4\, dx \rightarrow \frac{3}{4}\left(\frac{1}{2}x^4 - \frac{1}{5}x^5\right)\Big|_0^2 \rightarrow \frac{3}{4}\left(\frac{40}{5} - \frac{32}{5}\right)$$

$$= \frac{24}{20} \qquad Var(x) = \frac{6}{5} - 1 = \boxed{\frac{1}{5} = var(x)}$$

Figure 2: Homework 4 Page 2

c) $F_Y = P(Y \le t)$ $\quad Y = \sqrt{X}$

$F_Y = P(\sqrt{X} \le t)$

$F_Y = P(X \le t^2)$

$F_X(t) = P(X \le t) = \frac{(3-x)\,x^2}{4}$

$$\boxed{F_Y = \frac{(3-t^2)\,t^4}{4}}$$

d) $F_Y = \frac{3}{4}t^4 - \frac{1}{4}t^6$

$$\boxed{f_Y = 3t^3 - \frac{3}{2}t^5}$$

e) $\int_0^{\sqrt{2}} 3t^4 - \frac{3}{2}t^6 \, dt \rightarrow \frac{3}{5}t^5 - \frac{3}{14}t^7 \Big|_0^{\sqrt{2}}$

$3.394 - 2.424 = \boxed{.97 = E[Y]}$

$E[Y^2] - E[Y]^2$

$\int_0^{\sqrt{2}} t^2\left(3t^3 - \frac{3}{2}t^5\right) dt$ $\quad \int_0^{\sqrt{2}} 3t^5 - \frac{3}{2}t^7 \, dt \rightarrow \frac{1}{2}t^6 - \frac{3}{16}t^8 \Big|_0^{\sqrt{2}}$

$4 - 3 = 1$

$1 - .9409 = \boxed{.0591 = Var(Y)}$

Figure 3: Homework 4 Page 3

$$X = 17 \cdot (Y - 3) \qquad Y \sim N(3, 4^2)$$

4) a) $X = 17Y - 51$

mean $= 3$

var $= 4^2$

sd: 4

$$E[X] = E[17Y - 51] = 17E[Y] - 51$$

$E[Y] =$ mean

$$17(3) - 51 = \boxed{0 = E[Y]}$$

$$\text{Var}(17Y - 51) = 17^2 \text{ var}(Y) = 17^2 \cdot 4^2$$

$$\boxed{\text{Var}(X) = 4624}$$

b) $X = \sin(Y) \qquad Y \sim \text{Ber}(p)$

$$E[g(x)] = \sum g(a_i)\, p(a_i)$$

$$Y \begin{cases} 0 & 1-p \\ 1 & p \end{cases} \qquad X \begin{cases} \sin(0) & 1-p \\ \sin(1) & p \end{cases}$$

Figure 4: Homework 4 Page 4

c) $X \sim \text{pois}(\lambda)$ $\quad a_i = k$

$$E[x] = \sum k\left(\frac{\lambda^k}{k!}e^{-\lambda}\right) \qquad k = 0, 1, 2 \ldots$$

$$0 + \frac{\lambda}{1}e^{-\lambda} + \frac{2\lambda^2}{2!}e^{-\lambda} + \frac{3\lambda^3}{3!}e^{-\lambda} \ldots$$

$$= \lambda e^{-\lambda}\left(1 + \sum \frac{\lambda^{k-1}}{(k-1)!}\right) \qquad e^{\lambda} = 1 + \sum \frac{\lambda^{k-1}}{(k-1)!}$$

$$= \lambda e^{-\lambda} e^{\lambda} = \boxed{\lambda = E[x]}$$

$$Var(x) = E[x^2] - E[x]^2$$

$$k = 0, 1, 2 \ldots$$

$$E[x^2] = \sum k^2\left(\frac{\lambda^k}{k!}e^{-\lambda}\right)$$

$$0 + \frac{\lambda}{1}e^{-\lambda} + \frac{4\lambda^2}{2!}e^{-\lambda} + \frac{9\lambda^3}{3!}e^{-\lambda}$$

$$\lambda e^{-\lambda}\left(\sum \frac{(k-1)}{(k-1)!}\lambda^{k-1} + \sum \frac{1}{(k-1)!}\lambda^{k-1}\right)$$

$$\lambda e^{-\lambda}\left(\lambda \sum \frac{1}{(k-2)!}\lambda^{k-2} + \sum \frac{1}{(k-1)!}\lambda^{k-1}\right)$$

$$\lambda e^{-\lambda}\left(\lambda e^{\lambda} + e^{\lambda}\right)$$

$$\lambda^2 + \lambda$$

$$E[x^2] - E[x]^2 = \lambda^2 + \lambda - \lambda^2 = \boxed{\lambda = Var(x)}$$

Figure 5: Homework 4 Page 5

d) $X = U^2$ $\quad U \sim \text{unif}(-2, 5)$

$$E[X] = \int x \, f(x) \, du$$

$$E[X] = E[U^2] = \int_{-2}^{5} u^2 \frac{1}{5+2} \, du$$

$$\frac{1}{7} \int_{-2}^{5} u^2 \, du \rightarrow \frac{1}{7} \left(\frac{1}{3} u^3 \right) \Big|_{-2}^{5}$$

$$\frac{1}{21} (125 + 8) = \frac{133}{21} = \boxed{6.33 = E[X]}$$

$$\text{Var}(X) = E[X^2] - E[X]^2$$

$$E[X^2] = E[u^4] = \int_{-2}^{5} u^4 \frac{1}{7} \, du$$

$$\frac{1}{7} \int_{-2}^{5} u^4 \, du \rightarrow \frac{1}{35} \left(u^5 \right) \Big|_{-2}^{5}$$

$$\frac{3125}{35} + \frac{32}{35} = \frac{3157}{35} = E[X^2]$$

$$\frac{3157}{35} - \left(\frac{133}{21} \right)^2 = \boxed{50.09 = \text{Var}(X)}$$

Figure 6: Homework 4 Page 6

e) $X \sim \exp(\lambda)$ $\quad \int_0^\infty x \lambda e^{-\lambda x}\, dx$

$y = \lambda x$

$dy = \lambda$

$dx = \frac{dy}{\lambda}$

$$\frac{1}{\lambda}\int_0^\infty y e^{-y}\, dy$$

$$\frac{1}{\lambda}\int y e^{-y}\, dy \qquad u = y \quad v = e^{-y} \qquad du = 1 \quad dv = -e^{-y}$$

$$\frac{1}{\lambda}\left(-ye^{-y} - \int e^{-y}\, dy\right)$$

$$\frac{1}{\lambda}\left(-ye^{-y} + e^{-y}\right)\Big|_0^\infty$$

$$\boxed{\frac{1}{\lambda} = E[x]}$$

$$Var(x) = E[x^2] - E[x]^2$$

$$\int_0^\infty x^2 \lambda e^{-\lambda x}\, dx \qquad y = \lambda x \quad x = \frac{y}{\lambda} \qquad dx = \frac{dy}{\lambda}$$

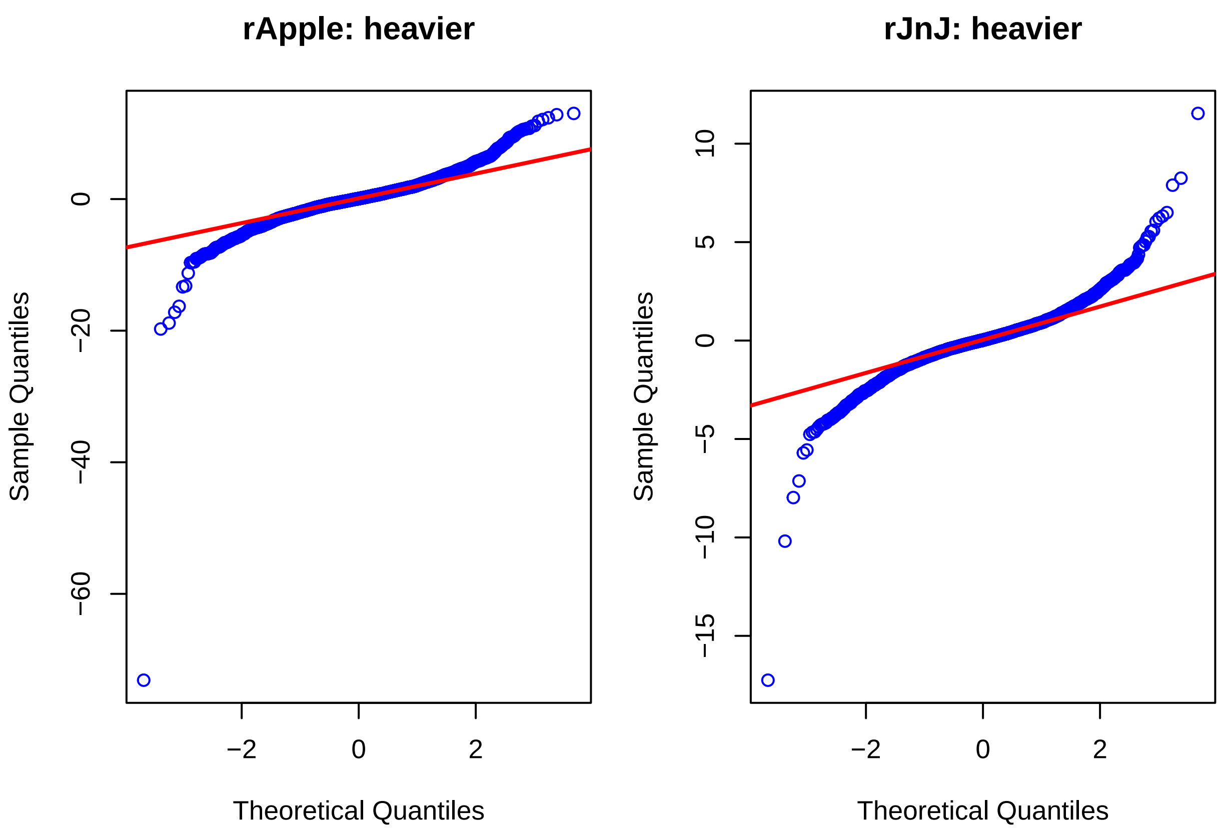
$$\frac{1}{\lambda^2}\int_0^\infty y^2 e^{-y}\, dy \qquad u = y^2 \quad v = e^{-y} \qquad du = 2y \quad dv = -e^{-y}$$

$$\frac{1}{\lambda^2}\left(-2e^{-y} - 2ye^{-y} - y^2e^{-y}\right)\Big|_0^\infty$$

$$\frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \boxed{\frac{1}{\lambda^2} = Var(x)}$$

Figure 7: Homework 4 Page 7

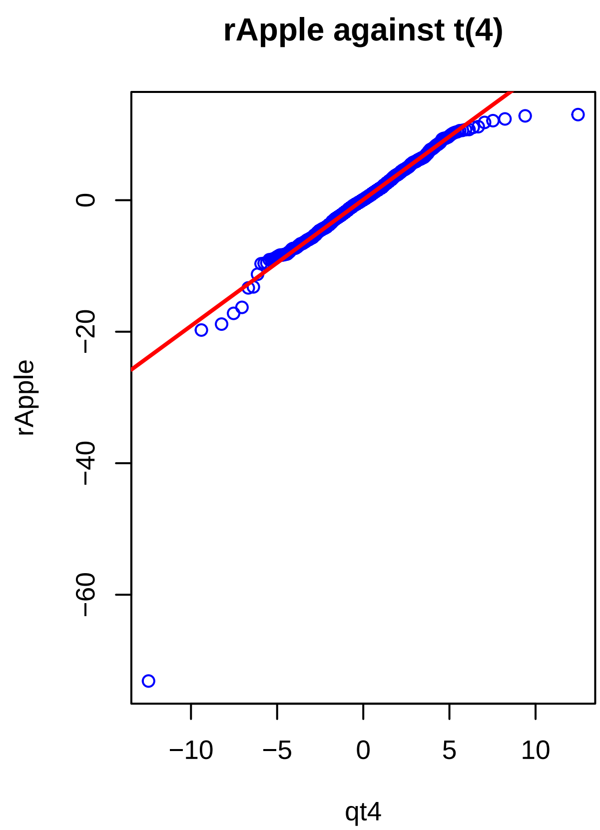
```r
pApple <- rev(apple[,7])
rApple <- diff(log(pApple))*100
pjnj <- rev(jnj[,7])
rjnj <- diff(log(pjnj))*100

par(mfrow=c(1,2), mar=c(4,4,4,1)+.1, cex=.8)

qqnorm(rApple, col="blue", main="")
qqline(rApple, lwd=2, col="red")
title("rApple: heavier")

qqnorm(rjnj, col="blue", main="")
qqline(rjnj, lwd=2, col="red")
title("rJnJ: heavier")
```

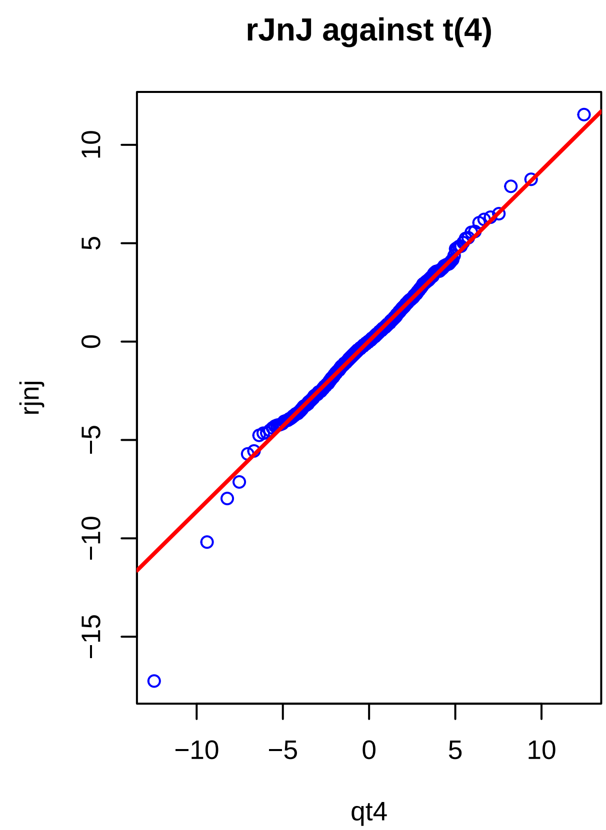


a). The log returns of Apple and JnJ do not follow a normal distribution because the tails are heavier.

```r
par(mfrow=c(1,2), mar=c(4,4,4,1)+.1, cex=.8)
x <- (1:length(rApple)-.5)/length(rApple)
qt4 <- qt(x, 4)
res <- qqplot(qt4, rApple, col="blue", main="")
x <- res$x
y <- res$y
abline(lsfit(x, y), col="red", lwd=2)
title("rApple against t(4)")
```

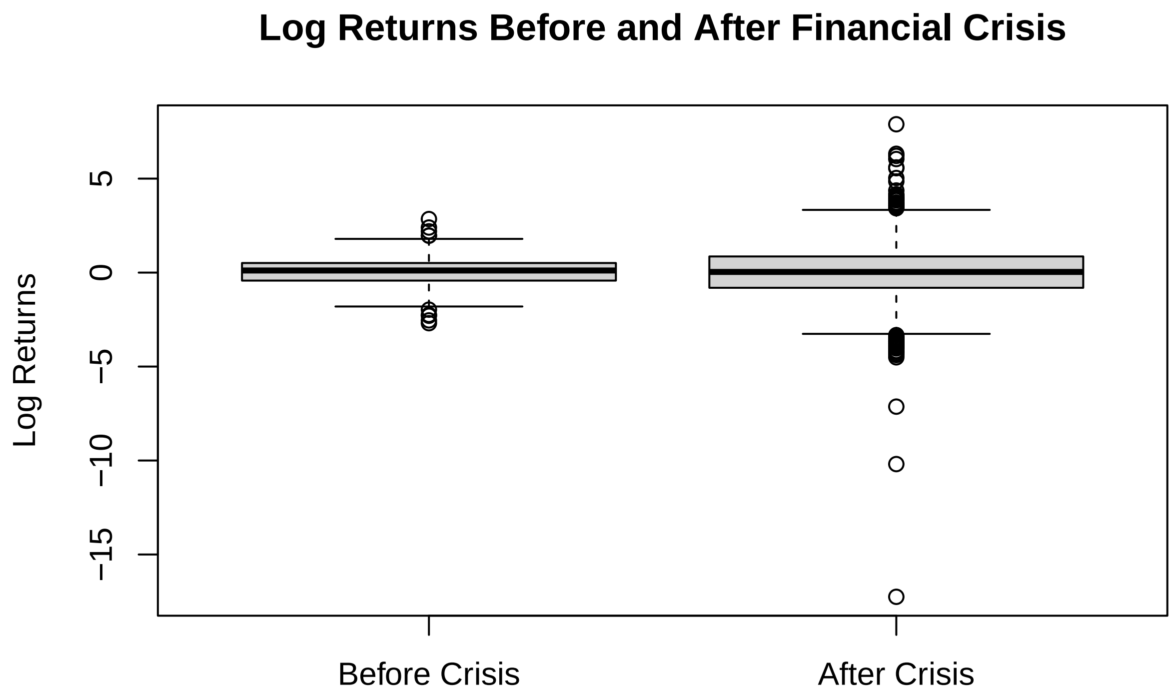
```r
x <- (1:length(rjnj)-.5)/length(rjnj)
qt4 <- qt(x, 4)
res <- qqplot(qt4, rjnj, col="blue", main="")
x <- res$x
y <- res$y
abline(lsfit(x, y), col="red", lwd=2)
title("rJnJ against t(4)")
```



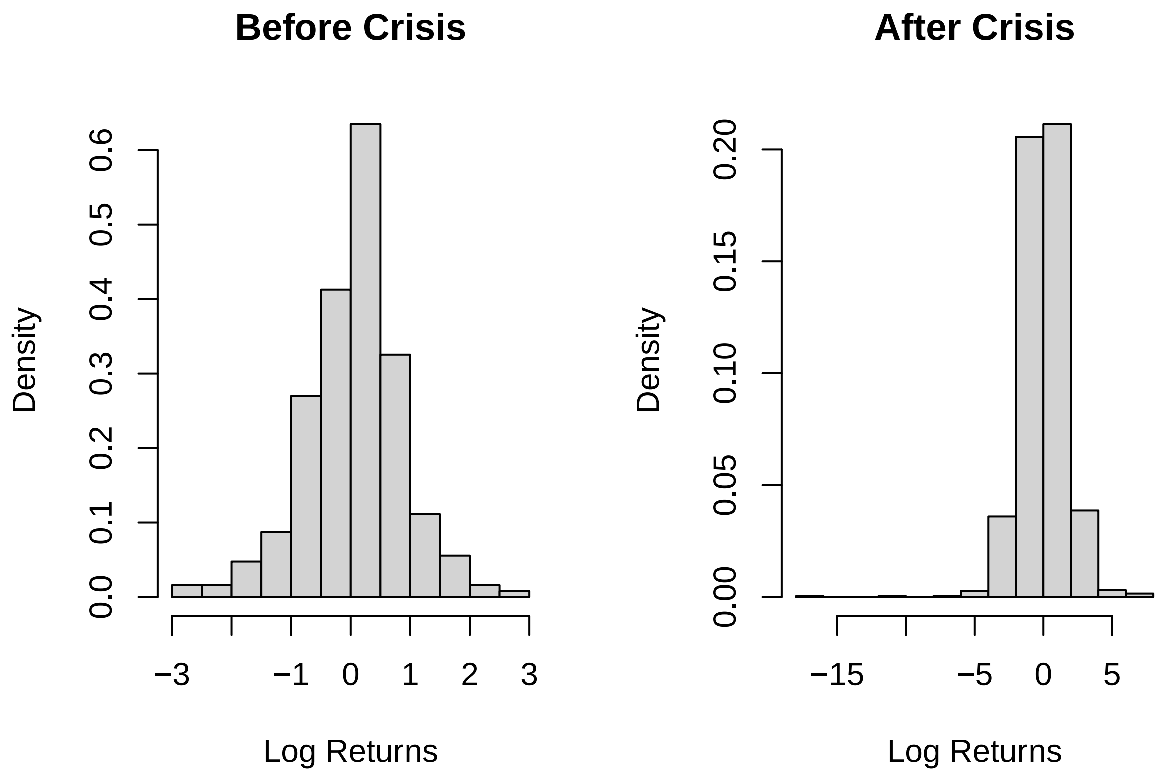


b). A t-distribution with 4 degrees of freedom seems to be a better fit because it reduces the heavy-tailedness of the log returns.

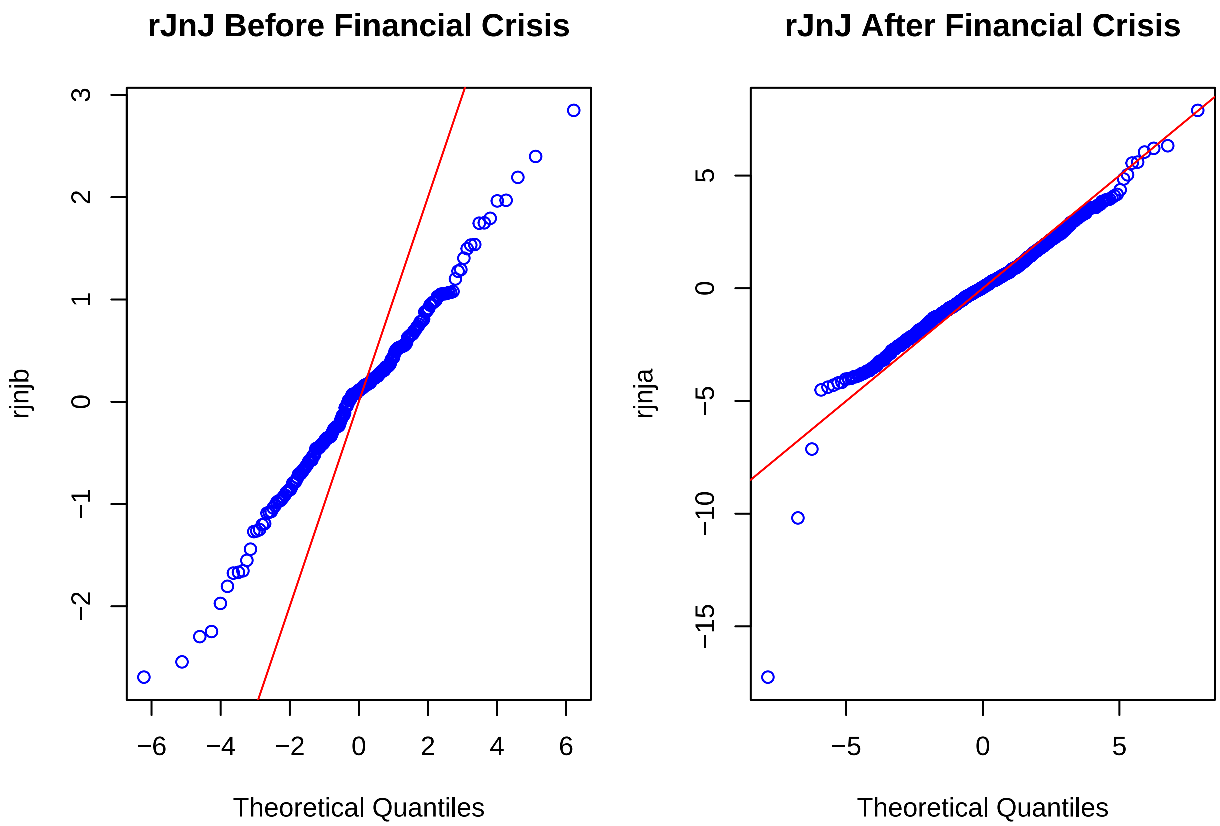
```r
rjnjb <- rjnj[2063:1812]
rjnja <- rjnj[1306:1]

boxplot(rjnjb, rjnja,
        names = c("Before Crisis", "After Crisis"),
        ylab = "Log Returns")
title("Log Returns Before and After Financial Crisis")
```




```
par(mfrow = c(1, 2))  # Create a 1x2 grid for plots

hist(rjnjb, breaks = 15, main = "Before Crisis",
     xlab = "Log Returns", freq=FALSE)

hist(rjnja, breaks = 15, main = "After Crisis",
     xlab = "Log Returns", freq=FALSE)
```

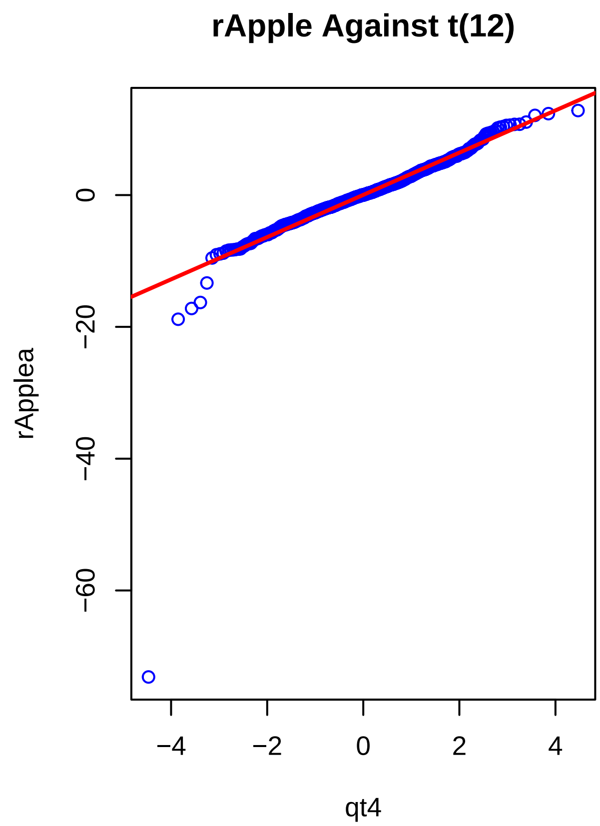


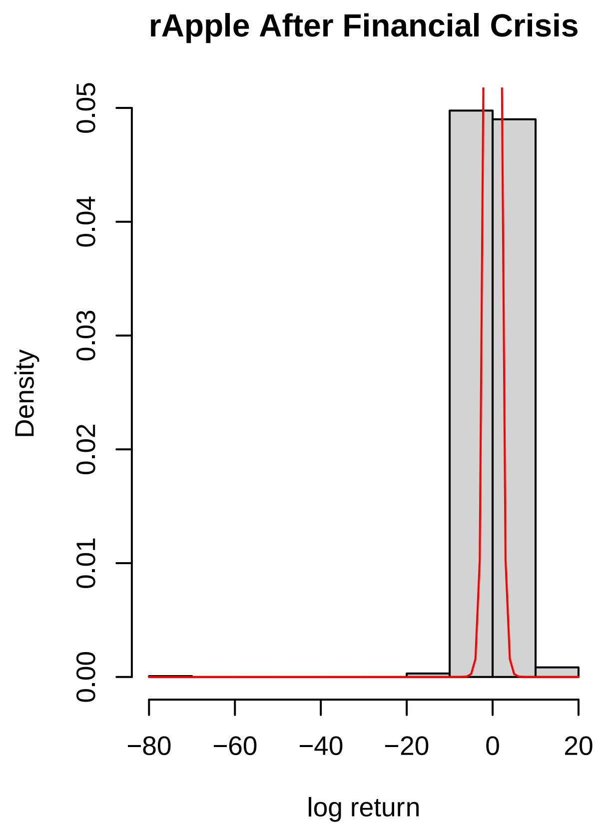

```r
par(mfrow=c(1,2), mar=c(4,4,4,1)+.1, cex=.8)
res <- qqplot(qlogis(ppoints(length(rjnjb))), rjnjb, col="blue", main="", xlab="Theoretical Quantiles")
abline(0, 1, col="red", main="")
title("rJnJ Before Financial Crisis")

res <- qqplot(qlogis(ppoints(length(rjnja))), rjnja, col="blue", main="", xlab="Theoretical Quantiles")
abline(0, 1, col="red", main="")
title("rJnJ After Financial Crisis")
```



c). Although there are more outliers after the financial crisis, it seems that the log returns after the financial crisis follows a normal distribution a lot better, as the tails aren't as light as before the financial crisis.

```
rApplea <- rApple[1306:1]

par(mfrow=c(1,2), mar=c(4,4,4,1)+.1, cex=.8)
x <- (1:length(rApplea)-.5)/length(rApplea)
qt4 <- qt(x, 12)
res <- qqplot(qt4, rApplea, col="blue", main="")
x <- res$x
y <- res$y
abline(lsfit(x, y), col="red", lwd=2)
title("rApple Against t(12)")

x <- (1:length(rApplea)-.5)/length(rApplea)
hist(rApplea, main="rApple After Financial Crisis", xlab="log return", ylab="Density", freq=FALSE)
curve(dt(x, df=12), add=TRUE, col='red' )
```





d). I believe 12 degrees of freedom is the best fit because it follows the QQ-Plot and histogram well.