FinalExam158

Henrique

May 6, 2019

```
n <- 30
coefs_age <- rep(0,1e4)
coefs_age2 <- rep(0,1e4)
for (i in 1:1e4) {
 Age=round(runif(n,min=18,max=70))
 Age2 <- Age^2
 HR <- 94-Age*0.5+Age2*0.0035+rnorm(n,sd=10)
 model <- lm(HR~Age+Age2)
 coefs_age[i] <- summary(model)$coefficients[2,1]
 coefs_age2[i] <- summary(model)$coefficients[3,1]

}
mean(coefs_age)

## [1] -0.4813071

mean(coefs_age2)

## [1] 0.003293108

summary(model)

##
## Call:
## lm(formula = HR ~ Age + Age2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.7835  -9.1653   0.2415   8.0607  22.3109
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 78.528993  17.734920   4.428 0.000142 ***
## Age         -0.152304   0.896598  -0.170 0.866380
## Age2         0.001963   0.010606   0.185 0.854507
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.01 on 27 degrees of freedom
## Multiple R-squared:  0.00147,    Adjusted R-squared:  -0.07249
## F-statistic: 0.01988 on 2 and 27 DF,  p-value: 0.9803
```

a)mean of coefs_age is -0.50 and coefs_age2 is 0.004, its seems right as the summary of one of the instances of the model shows as coef for age being -.53 and for age2 to be -.005 which is pretty close to the mean.

```
n <- 30
ageee <- rep(0,1e4)
HRrr <- rep(0,1e4)
set.seed(104)
for (i in 1:1e4) {
 Age=round(runif(n,min=18,max=70))
 Age2 <- Age^2
 HR <- 94-Age*0.5+Age2*0.0035+rnorm(n,sd=10)
 model <- lm(HR~Age+Age2)
#ageee[i]<- mean(Age)
#HRrr[i]<- mean(HR)

}
plot(Age,HR, xlab='Age', ylab = 'Heart Rate', main = 'Heart Rate vs Age',
cex=0.9)
```

c)

```
n <- 30
p1 <- rep(0,1e4)
p2 <- rep(0,1e4)
for (i in 1:1e4) {
 Age=round(runif(n,min=18,max=70))
 Age2 <- Age^2
```

```
 HR <- 94-Age*0.5+Age2*0.0035+rnorm(n,sd=10)
 model <- lm(HR~Age+Age2)
 p1[i] <- summary(model)$coefficients[2,4]
 p2[i] <- summary(model)$coefficients[3,4]
}
sum(p1<0.05)/length(p1)

## [1] 0.0892

sum(p2<0.05)/length(p2)

## [1] 0.065
```

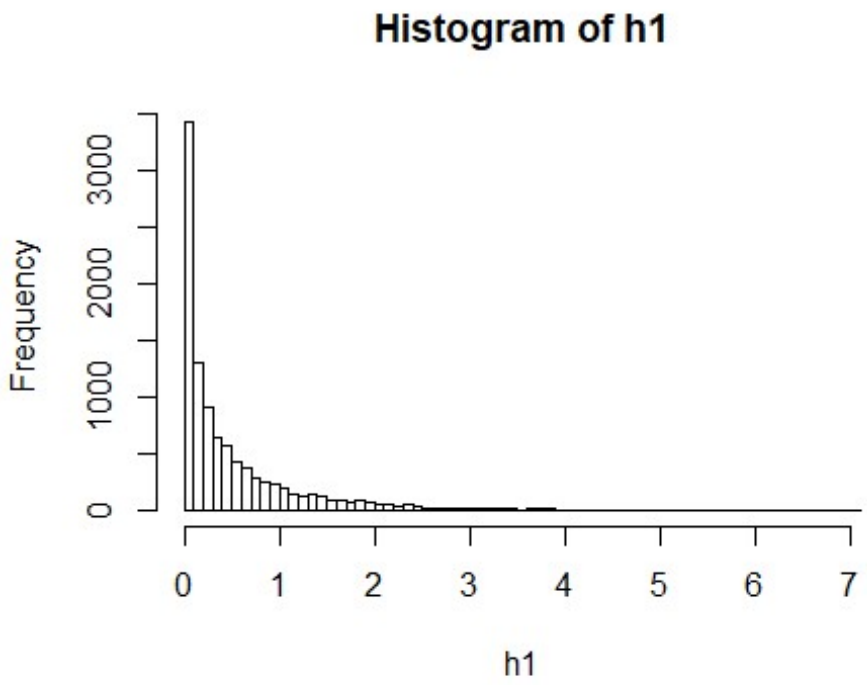
Power of age is 0.09 and the power of age^2 is 0.06

```
n <- 30
p1 <- rep(0,1e4)
for (i in 1:1e4) {
 Age=round(runif(n,min=18,max=70))
 Age2 <- Age^2
 HR <- 94-Age*0.5+Age*0.0035+rnorm(n,sd=10)
 model <- lm(HR~Age)
 p1[i] <- summary(model)$coefficients[2,4]
}
sum(p1<0.05)/length(p1)

## [1] 0.9632

summary(model)

##
## Call:
## lm(formula = HR ~ Age)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -29.706  -7.511  -1.041   9.088  20.993
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  92.4741     6.4524   14.33 2.03e-14 ***
## Age          -0.4518     0.1373   -3.29  0.00271 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.27 on 28 degrees of freedom
## Multiple R-squared:  0.2788, Adjusted R-squared:  0.253
## F-statistic: 10.82 on 1 and 28 DF,  p-value: 0.00271
```

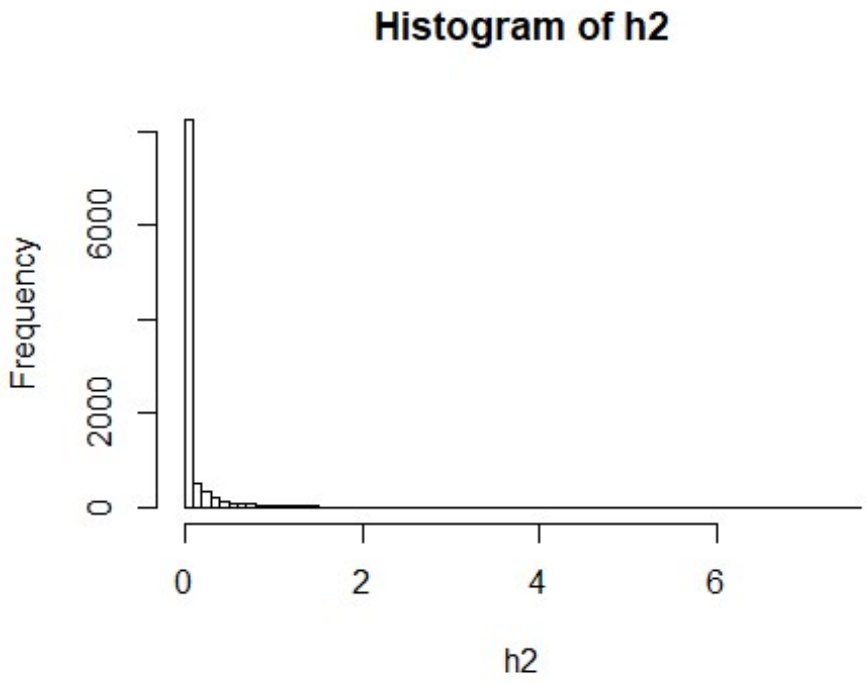
It increased the power by a lot which would mean that the relation between HR and age it is not quadratic.

2)a)

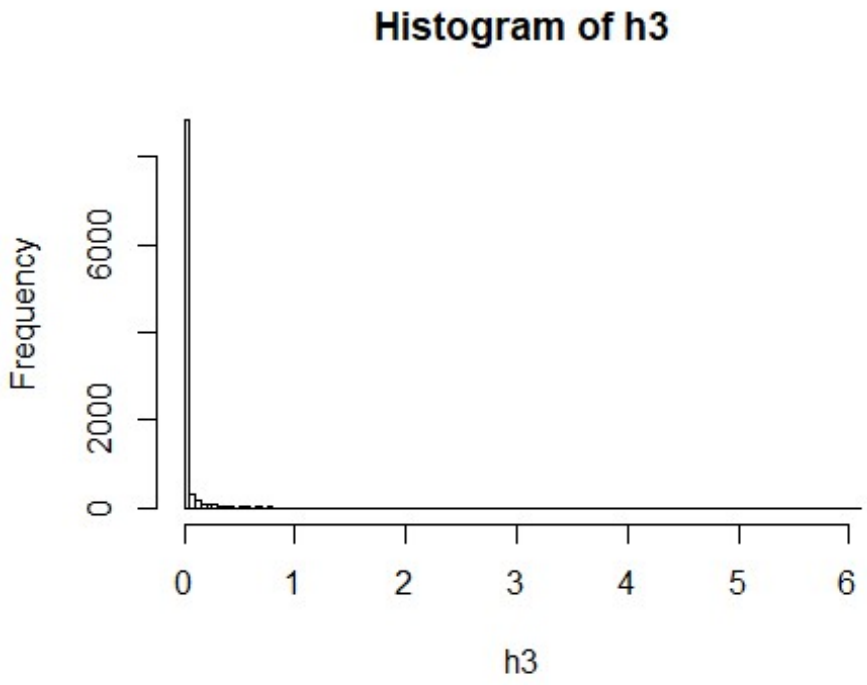
```
h1 <- rgamma(10000, shape = 0.5)
h2 <- rgamma(10000, shape =0.1)
h3 <- rgamma(10000, shape =0.05)

hist(h1,breaks=100)
```




```
hist(h2,breaks=100)
```

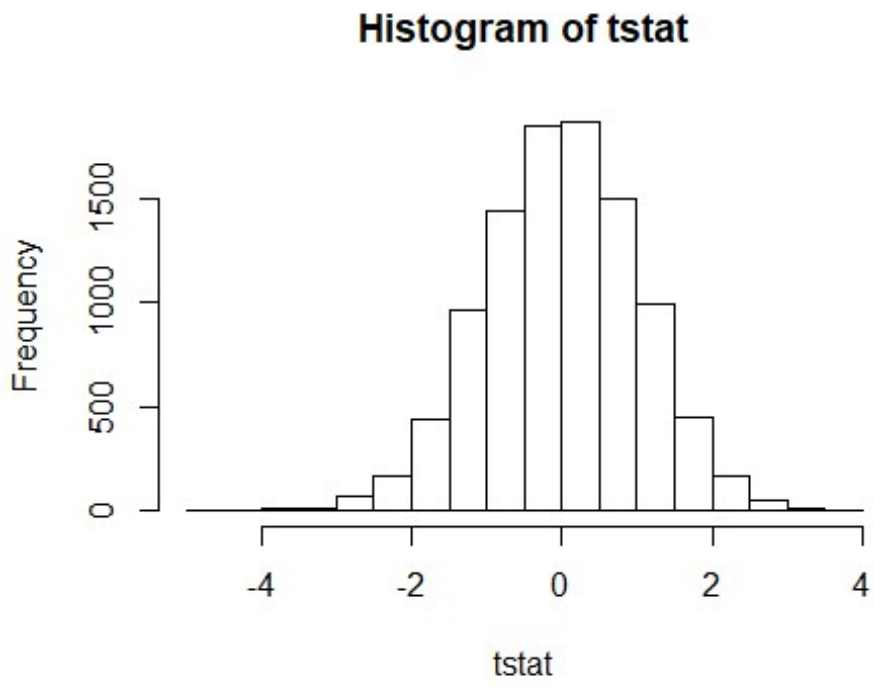



```
hist(h3,breaks=100)
```



As shape gets smaller the histogram gets more skewed to the right.

```r
tstat <- rep(1e4)
pvav <- rep(1e4)

for (i in 1:1e4) {
sim <- rgamma(30, shape = 1)
sim1 <- rgamma(30, shape = 1)

t <- t.test(sim,sim1)
tstat[i] <- t$statistic
pvav[i] <- t$p.value
}
hist(tstat)
```

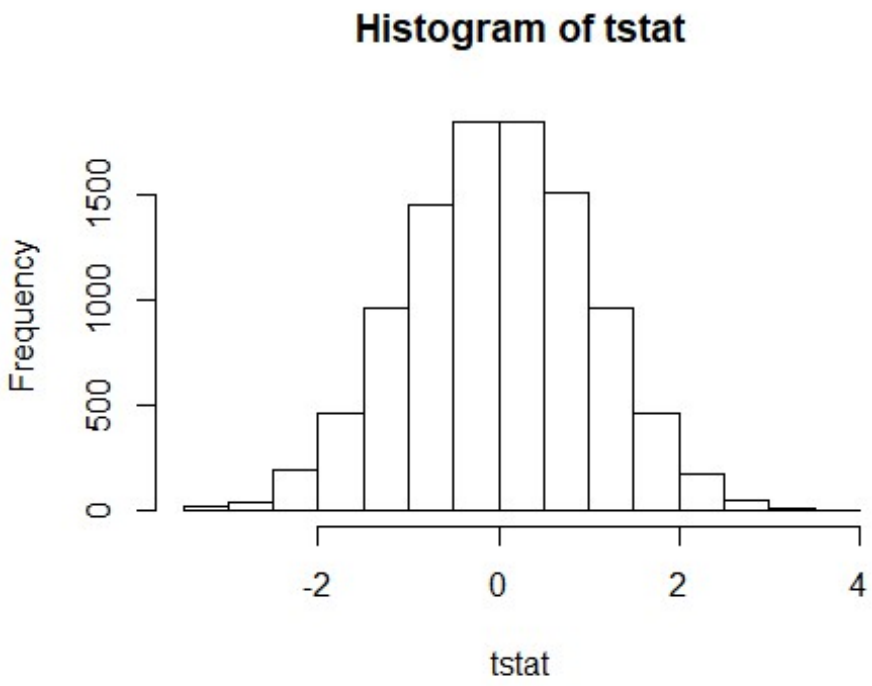



```r
sum(pvav<0.05)/length(pvav)
```

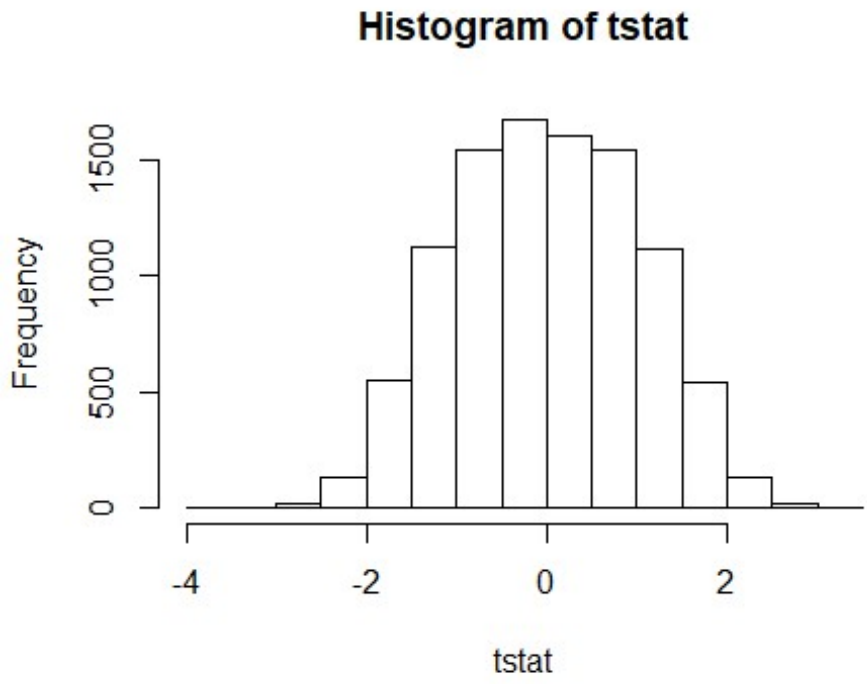
```
## [1] 0.046
```

```r
tstat <- rep(1e4)
pvav <- rep(1e4)

for (i in 1:1e4) {
sim <- rgamma(30, shape = 0.5)
sim1 <- rgamma(30, shape = 0.5)

t <- t.test(sim,sim1)
tstat[i] <- t$statistic
pvav[i] <- t$p.value
```

```
}
hist(tstat)
```



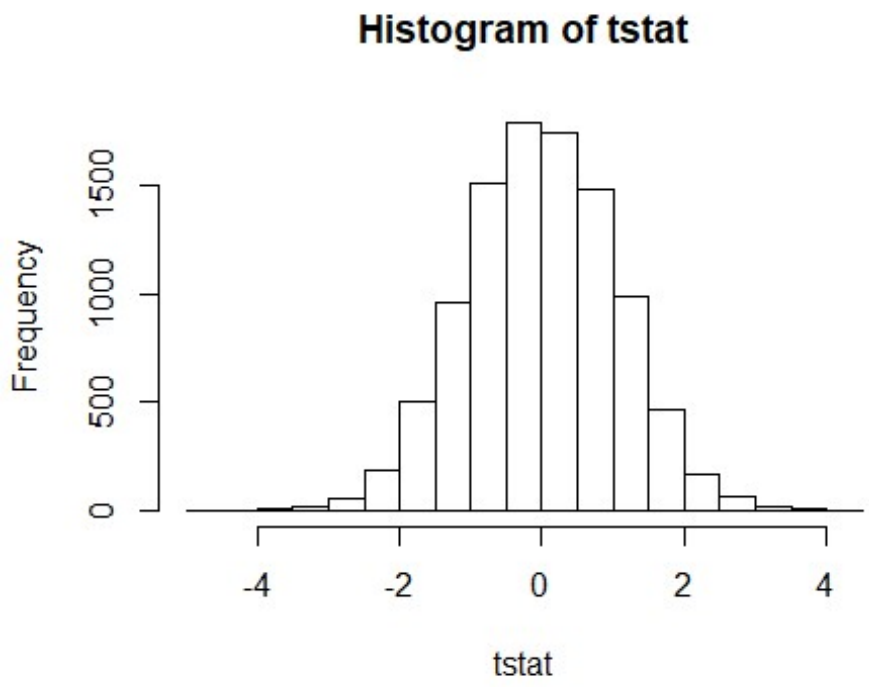

```
sum(pvav<0.05)/length(pvav)

## [1] 0.0466

tstat <- rep(1e4)
pvav <- rep(1e4)

for (i in 1:1e4) {
sim <- rgamma(30, shape = 0.1)
sim1 <- rgamma(30, shape = 0.1)

t <- t.test(sim,sim1)
tstat[i] <- t$statistic
pvav[i] <- t$p.value
}
hist(tstat)
```



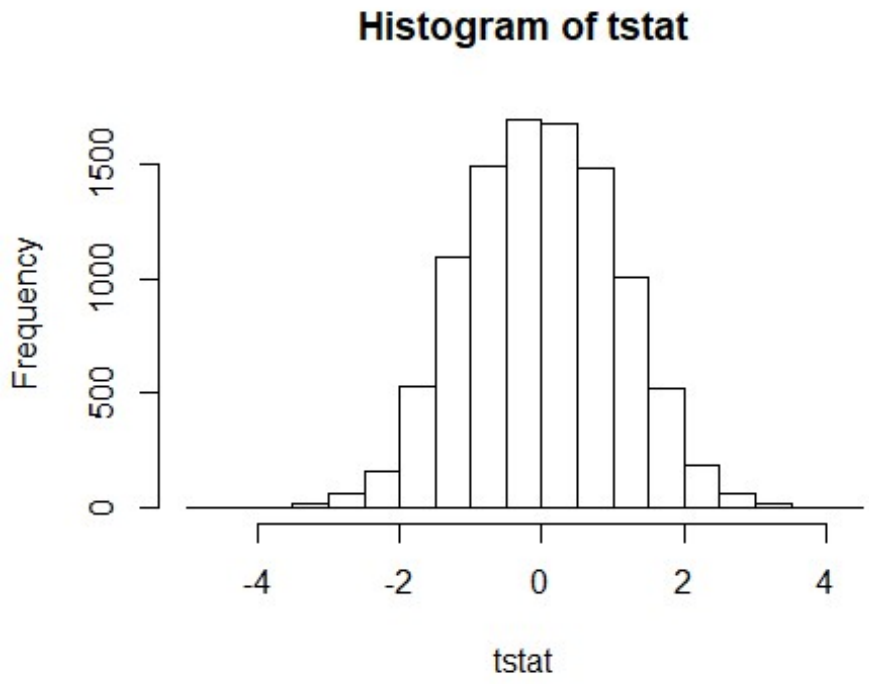

```
sum(pvav<0.05)/length(pvav)

## [1] 0.0262
```

the distribution is less normal as the shape decreases, also as the shape gets smaller the proportion of pvalues under 0.05 also gets smaller.

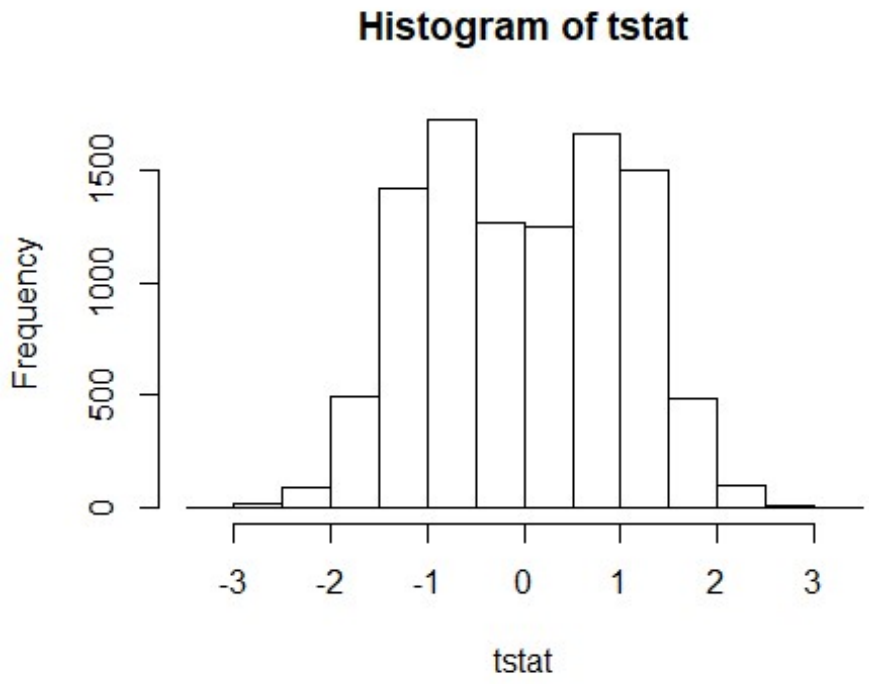
```
tstat <- rep(1e4)
pvav <- rep(1e4)

for (i in 1:1e4) {
sim <- rgamma(10, shape = 1)
sim1 <- rgamma(10, shape = 1)

t <- t.test(sim,sim1)
tstat[i] <- t$statistic
pvav[i] <- t$p.value
}
hist(tstat)
```




```r
sum(pvav<0.05)/length(pvav)

## [1] 0.0381

tstat <- rep(1e4)
pvav <- rep(1e4)

for (i in 1:1e4) {
sim <- rgamma(10, shape = 0.5)
sim1 <- rgamma(10, shape = 0.5)

t <- t.test(sim,sim1)
tstat[i] <- t$statistic
pvav[i] <- t$p.value
}
hist(tstat)
```




```
sum(pvav<0.05)/length(pvav)

## [1] 0.032

tstat <- rep(1e4)
pvav <- rep(1e4)

for (i in 1:1e4) {
sim <- rgamma(10, shape = 0.1)
sim1 <- rgamma(10, shape = 0.1)

t <- t.test(sim,sim1)
tstat[i] <- t$statistic
pvav[i] <- t$p.value
}
hist(tstat)
```




```
sum(pvav<0.05)/length(pvav)
```

```
## [1] 0.0074
```

yes, as the shape and the n go down the distribution of the mean is less and less normal.

3a)

```
ferret = read.csv("Ferret_Vaccine.csv", header = TRUE)
attach(ferret)
mean(Temperature)
```

```
## [1] 101.7395
```

```
sd(Temperature)
```

```
## [1] 0.9505654
```

```
mean(Weight)
```

```
## [1] 1145.234
```

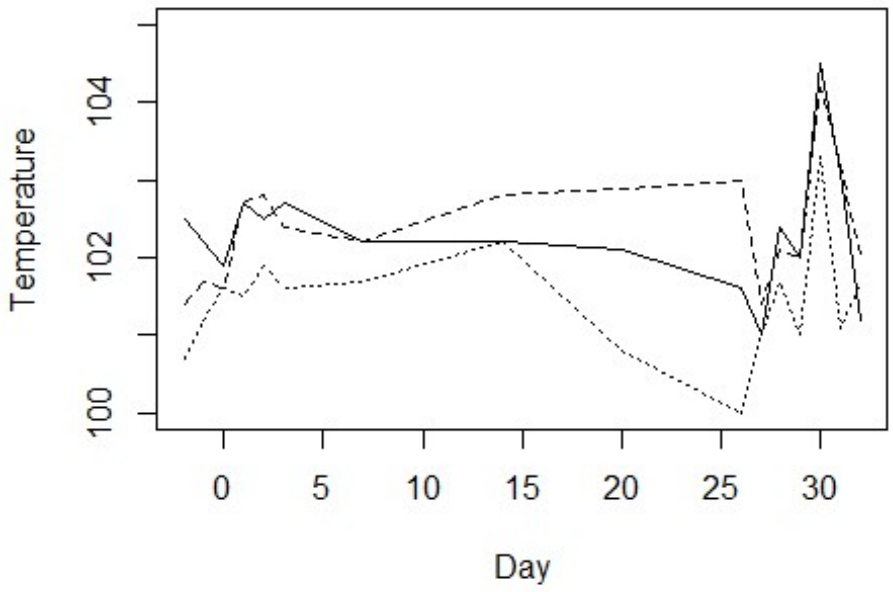
```
sd(Weight)
```

```
## [1] 148.4198
```

```
detach(ferret)
```

Temperature mean = 101.7395 Temperature SD= 0.9505654 Weight mean = 1145.234 Weight SD = 148.4198

```
attach(ferret)

plot(Day[Ferret.ID==574],Temperature[Ferret.ID==574],type="l",xlab='Day',ylab
= 'Temperature', ylim=c(100,105))
lines(Day[Ferret.ID==546],Temperature[Ferret.ID==546],lty = 2)
lines(Day[Ferret.ID==548],Temperature[Ferret.ID==548], lty = 3)
```




```
detach(ferret)
```