homework5

3170102187　　孙晨旭

1.

```
percentile_ratio_discrepancies <- function(a,P99,P99.5,P99.9){
  re <- (((P99/P99.9)^(-a+1)-10)^2 +((P99.5/P99.9)^(-a+1)-5)^2+((P99/P99.5)^(-a+1)-2)^2
  return(re)
}
```

```
percentile_ratio_discrepancies(a = 2,P99 = 1e6,P99.5 = 2e6,P99.9 = 1e7)
```

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

2.

```
exponent.multi_ratios_est <- function(A){
#  A = c(P99,P99.5,P99.9)
  init = 1-log(10)/log(A[1]/A[3])
  T<-nlm(percentile_ratio_discrepancies,init,A[1],A[2],A[3])
  return(T$estimate)
}
```

```
exponent.multi_ratios_est(c(1e6,2e6,1e7))
```

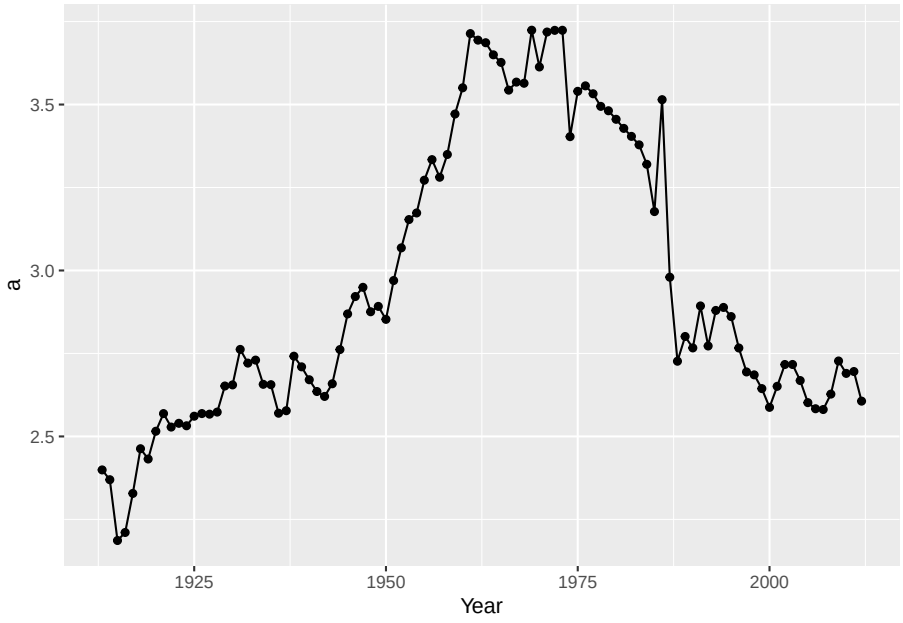
```
## [1] 2
```

3.

```
wtid_1 <- wtid %>% select(`P99 income threshold`,
                          `P99.5 income threshold`,
                          `P99.9 income threshold`)
a_est <- function(data = wtid_1){
  re<-apply(data,1,exponent.multi_ratios_est)
  return(re)
}
```

```
wtid <- cbind(wtid,a = a_est())
```

```
wtid %>% ggplot(aes(x = Year, y = a))+
  geom_line()+
  geom_point()
```



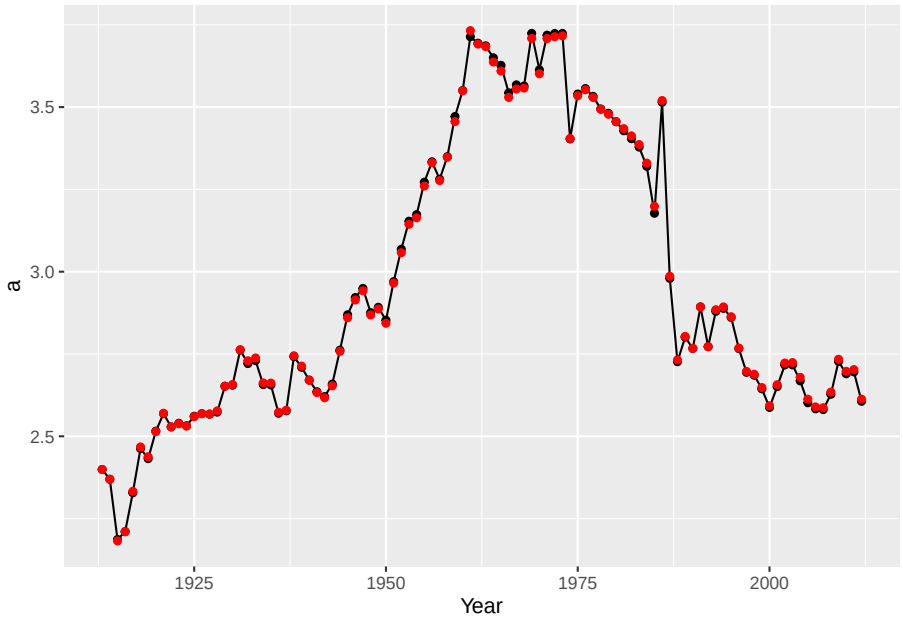
4.

```
wtid <- wtid %>% mutate(model_a = 1-log(10)/log(`P99 income threshold`/`P99.9 income th
summary(wtid$a-wtid$model_a)
```

```
##       Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
```

```
## -0.0210058 -0.0052339 -0.0003915  0.0001066  0.0041078  0.0173385
```

```
wtid %>% ggplot(aes(x = Year, y = a))+
  geom_line()+
  geom_point()+
  geom_point(aes(x = Year,y = model_a),col = 'red')
```



从这些数据来看，两个估计比较接近，说明 $a = 1 - \frac{\log(10)}{\log(P99/P99.9)}$ 是第一问中优化问题的一个比较好的估计。