Homework 5: Pareto and Kuznets on the Grand Tour

Shen Dingtao 3170104764

Load the data set

1.

```
percentile_ratio_discrepancies<-function(params=c(P99,P99.5,P99.9,a)){
  result<-((params[1]/params[3])^(1-params[4])-10)^2+
    ((params[2]/params[3])^(1-params[4])-5)^2+
    ((params[1]/params[2])^(1-params[4])-2)^2
  return(result)
}
```

Check : `P99=1e6`, `P99.5=2e6`, `P99.9=1e7` and `a=2`

```
percentile_ratio_discrepancies(c(1e6,2e6,1e7,2))
```

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

2.

```
exponent.multi_ratios_est<-function(params=c(P99,P99.5,P99.9)){
  a<-1-log(10)/(log(params[1])-log(params[3]))
  par<-c(params,a)
  result<-nlm(percentile_ratio_discrepancies,par)$estimate[4]
  return(result)
}
```

Check: `P99=1e6`, `P99.5=2e6` , `P99.9=1e7`

```
exponent.multi_ratios_est(c(1e+06, 2e+06, 1e+07))
```

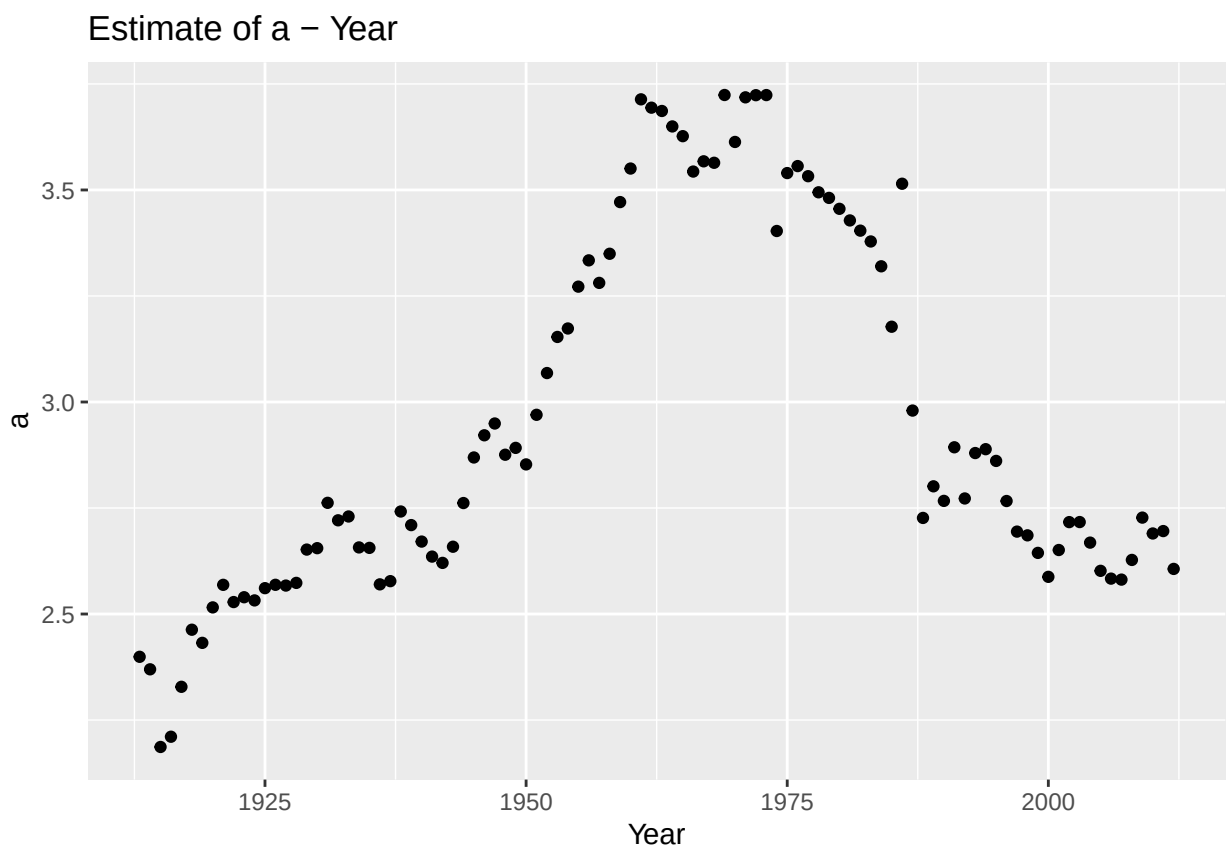
```
## [1] 2
```

3. Write a function which uses `exponent.multi_ratios_est` to estimate a for the US for every year from 1913 to 2012.

```
a_Esti_for_year<-function(){
  data<-wtid %>% select(-year)
  a_Esti<-apply(wtid %>% select(-year),1,exponent.multi_ratios_est)
  return(a_Esti)
}
# Estimate a from 1913 to 2012
a_E_Y<-a_Esti_for_year()
```

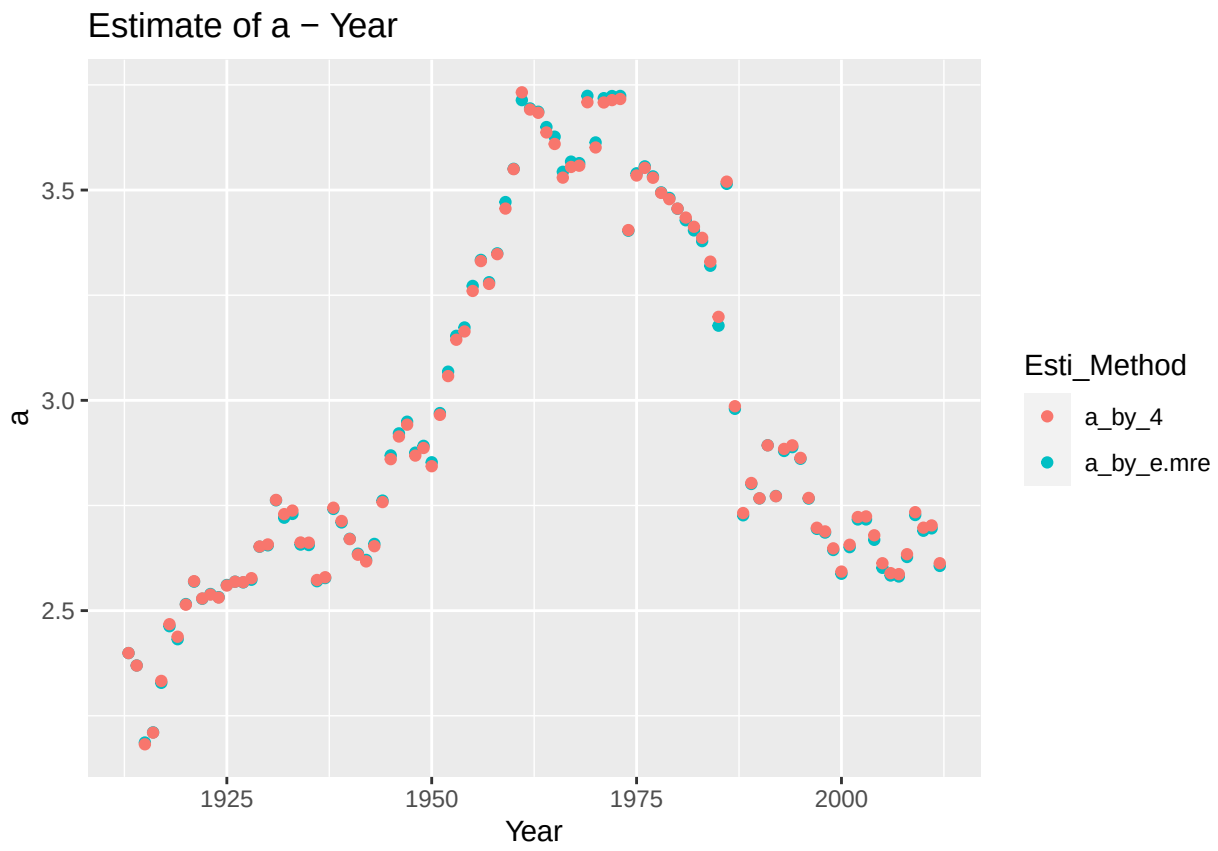
Plot the estimates:

```
ggplot()+
  geom_point(aes(x=wtid$year,y=a_E_Y))+
  labs(x="Year",y="a",title="Estimate of a - Year")
```



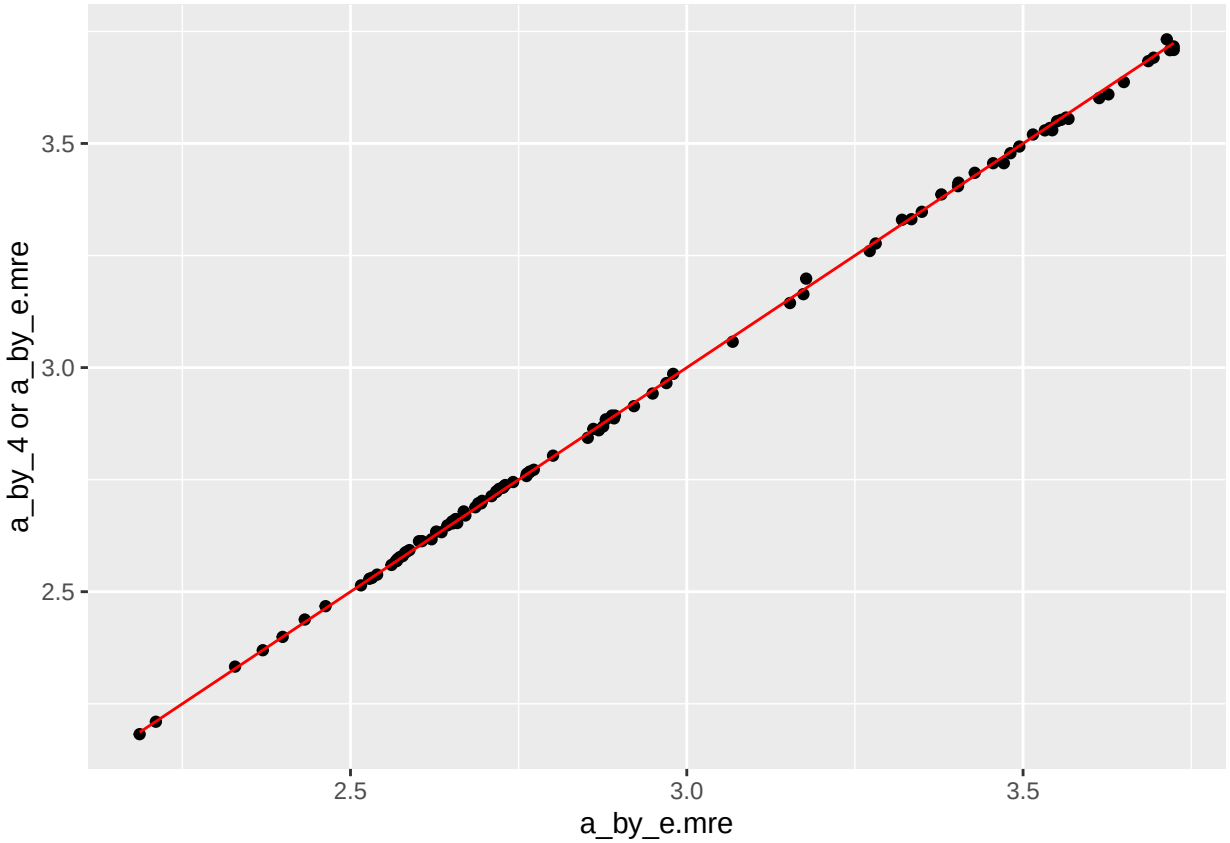
4.

```
a_E_Y2<-1 - log(10)/(log(wtid$p99)-log(wtid$p99.9))
df <- data.frame(year=wtid$year,a_by_e.mre = a_E_Y, a_by_4 = a_E_Y2)
df <-gather(df,Esti_Method,esti,-year)
ggplot(df) +
  geom_point(aes(x = year, y = esti, color = Esti_Method)) +
  labs(x="Year",y="a",title="Estimate of a - Year")
```



We can see that plots of the two estimates are very close to each other. We can plot `a_by_4-a_by_e.mre` as follows to see the comparison more clearly:

```
ggplot()+
  geom_point(aes(x=a_E_Y,y=a_E_Y2),color="black")+
  geom_line(aes(x=a_E_Y,y=a_E_Y),color="red")+
  labs(x="a_by_e.mre",y="a_by_4 or a_by_e.mre")
```


3.5
3.0
2.5
a_by_4 or a_by_e.mre
2.5
3.0
3.5
a_by_e.mre