Progress Report

Victor Huang

November 29, 2020

Overview:

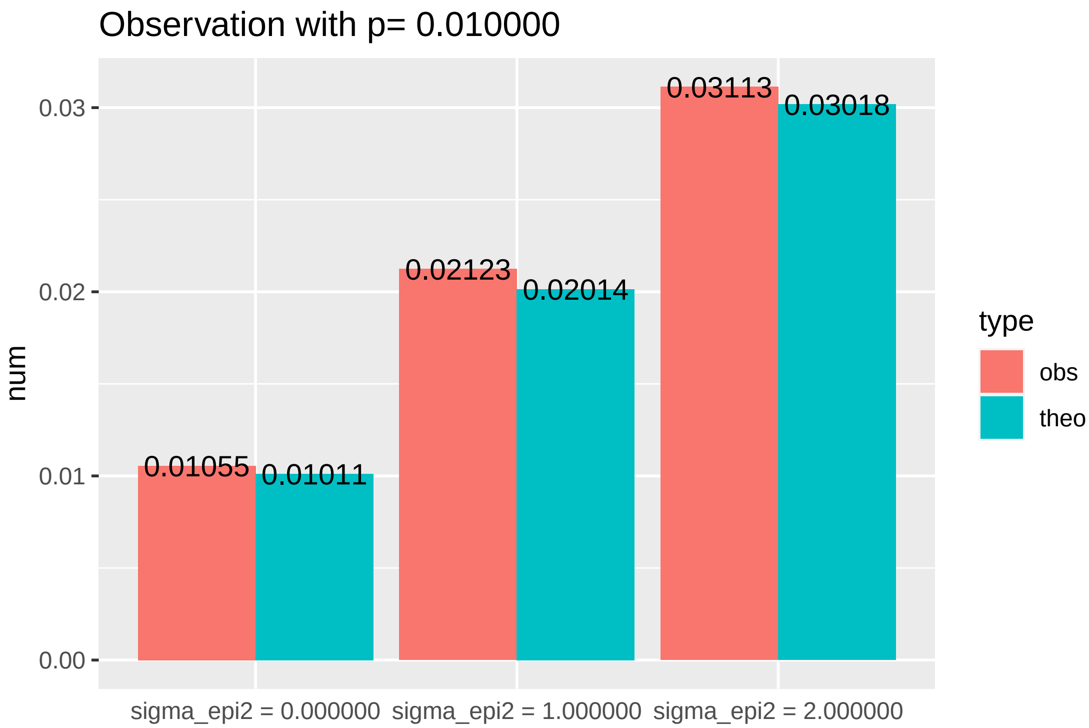
Concerned variables and underlying Monte-Carlo model:

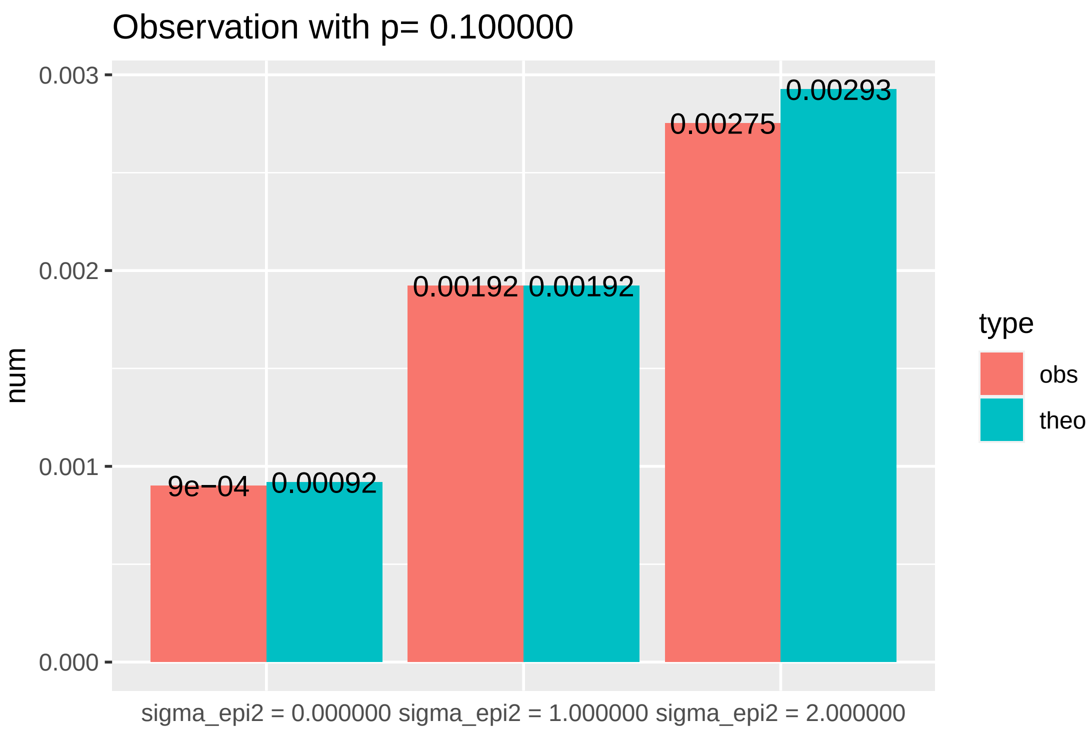
```
nSims: Number of simulations to run for each groups of parameter.
N: Size of population.
p: Probablity an observation is sampled.
sigma_e2: Variance of Model error e.
sigma_epi2: variance of unpredicted error epislon.
yi: Dependent variable without unpredicted error.
xi: Independent variable.
cap_y: Dependent variable with unpredicted error.

Simulation2: A function that randomly draws samples from population given. Each observation \\
has probability p of being samapled. It repeats this operation nSims times and will output a \\
vector of beta, the empirical value for an OLS fitting model based on cap_y and xi.
```

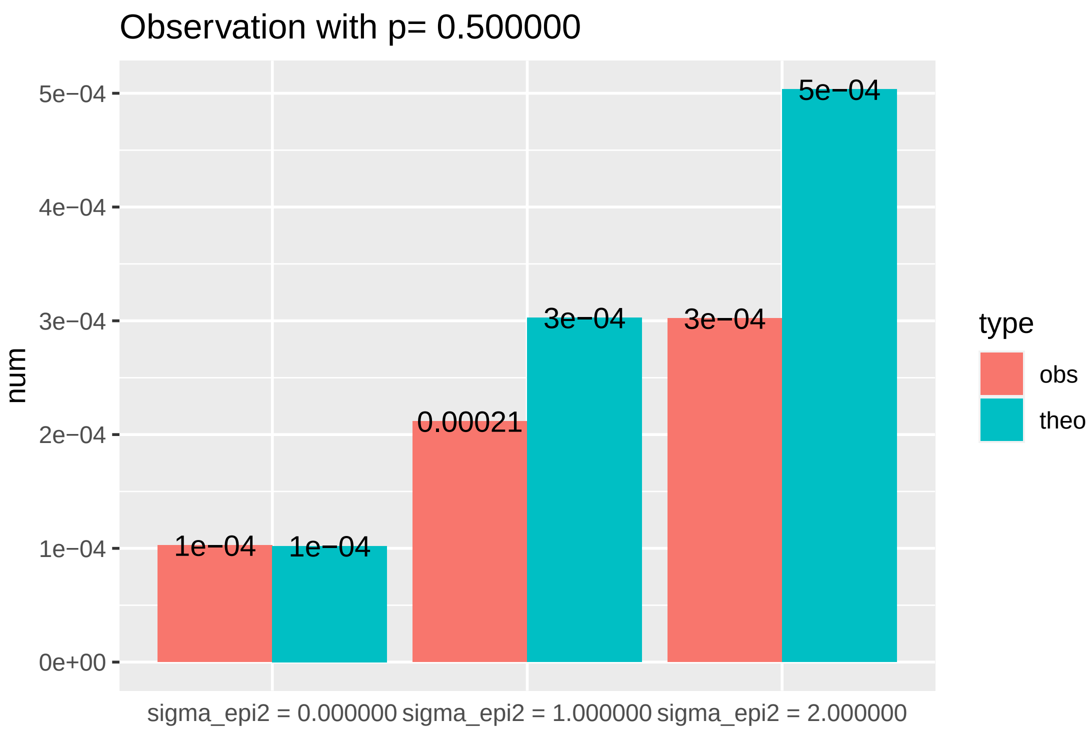
```r
# setting parameters
sigma_e2<-1
N<-10000
beta<-2
nSims<-10000

p<-c(0.01,0.1,0.5,0.9,0.99)
sigma_epi2<-c(0,1,2)

x_para<-set_para(0,1)
e_para<-set_para(0,sigma_e2)

# generating population
xi<-pop_gen(x_para,N,'normal')
ei<-pop_gen(e_para,N,'normal')

yi<-xi*beta+ei



result<-NULL

# for each group of simulations, generate one population with
```

```r
# independent error epislon.

for (j in c(1:length(sigma_epi2))){
    epi_para<- set_para(0,sigma_epi2[j])
    epi<-pop_gen(epi_para,N,'normal')

    cap_y<-yi+epi

# sample from the population with probability p and return the
# variance of the beta.
  for (i in c(1:length(p))){
    start.time.small<-Sys.time()
    var_beta<-var(simulation2(cap_y,xi,p[i],nSims))
    theo_var<-sigma_epi2[j]/(sum(xi**2))+((1-p[i])/p[i])*
      ((sum(xi**2*ei**2))/((sum(xi**2))**2)+sigma_epi2[j]/(sum(xi**2)))

    end.time.small<-Sys.time()

# rearranging the data for presentation and making graph
    out<-cbind(obs=var_beta,theo=theo_var,prob=sprintf('p = %f ',p[i]),
               sigma=sprintf('sigma_epi2 = %f',sigma_epi2[j]),
               time=end.time.small-start.time.small)
    result<-rbind(result,out)
  }
}
```

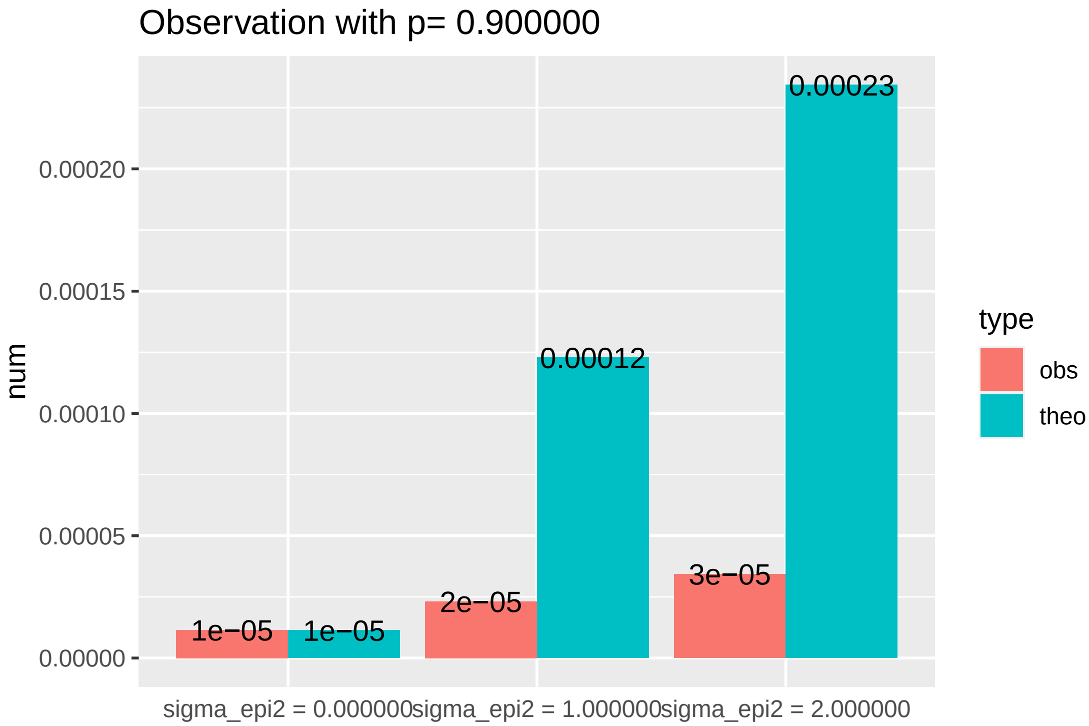
```r
result1<-as.data.frame(result) %>% gather(type,num,obs:theo) %>% type_convert()

for (i in (c(1:length(p)))){
a<-ggplot(result1 %>% filter(prob==sprintf('p = %f',p[i])))+
    geom_col(aes(sigma,num,fill=type),position='dodge',width = 0.9)+
  geom_text(aes(sigma,num,label=round(as.numeric(num),5),group=type),
            position = position_dodge(width = 0.9))+
  theme(axis.title.x=element_blank(),
        axis.ticks.x=element_blank(),
        plot.margin = unit(c(1,1,1,1), "cm"))+
  ggtitle(sprintf('Observation with p= %f',p[i]))
plot(a)
}
```


Observation with p= 0.010000
num
0.03
0.02
0.01
0.00
0.01055
0.01011
0.02123
0.02014
0.03113
0.03018
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo

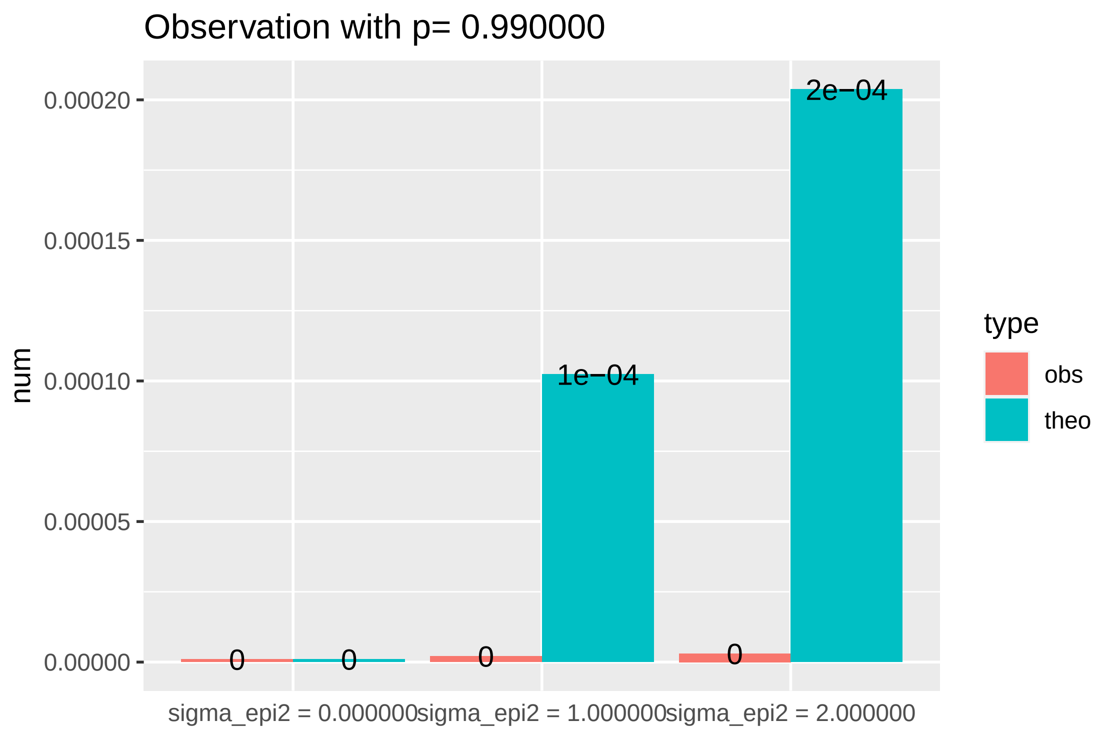

Observation with p= 0.100000
0.003
0.002
0.001
0.000
num
9e−04
0.00092
0.00192
0.00192
0.00275
0.00293
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo



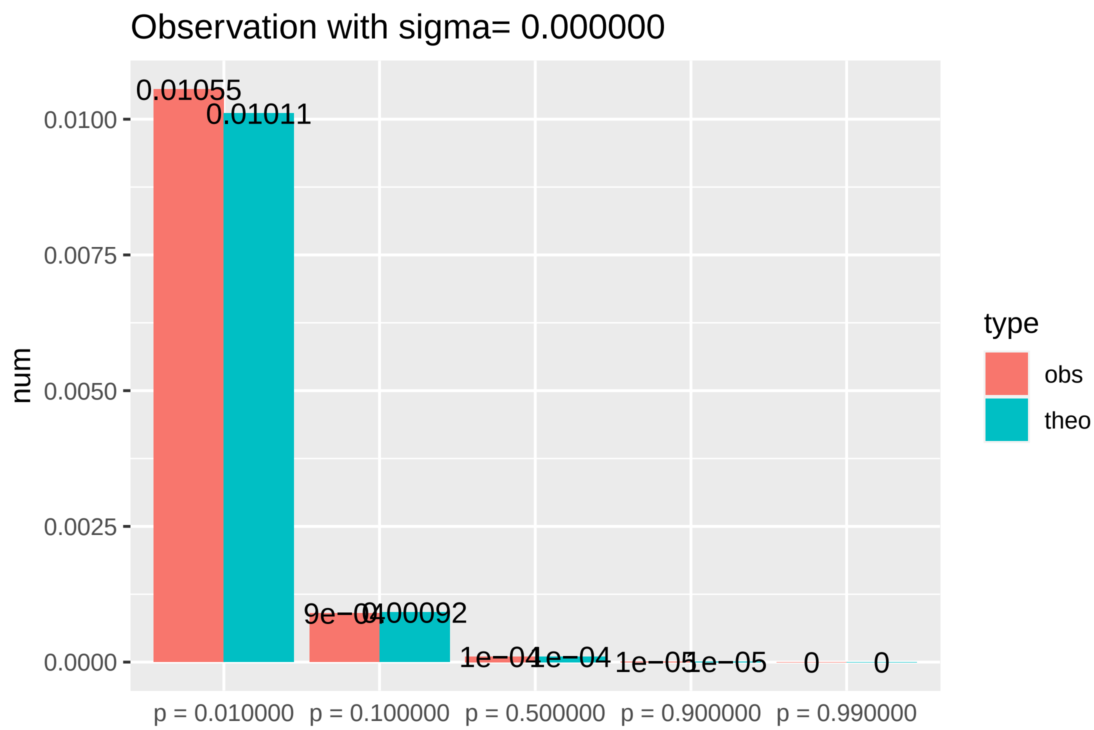
Observation with p= 0.500000
5e−04
4e−04
3e−04
2e−04
1e−04
0e+00
num
1e−04
1e−04
0.00021
3e−04
3e−04
5e−04
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo

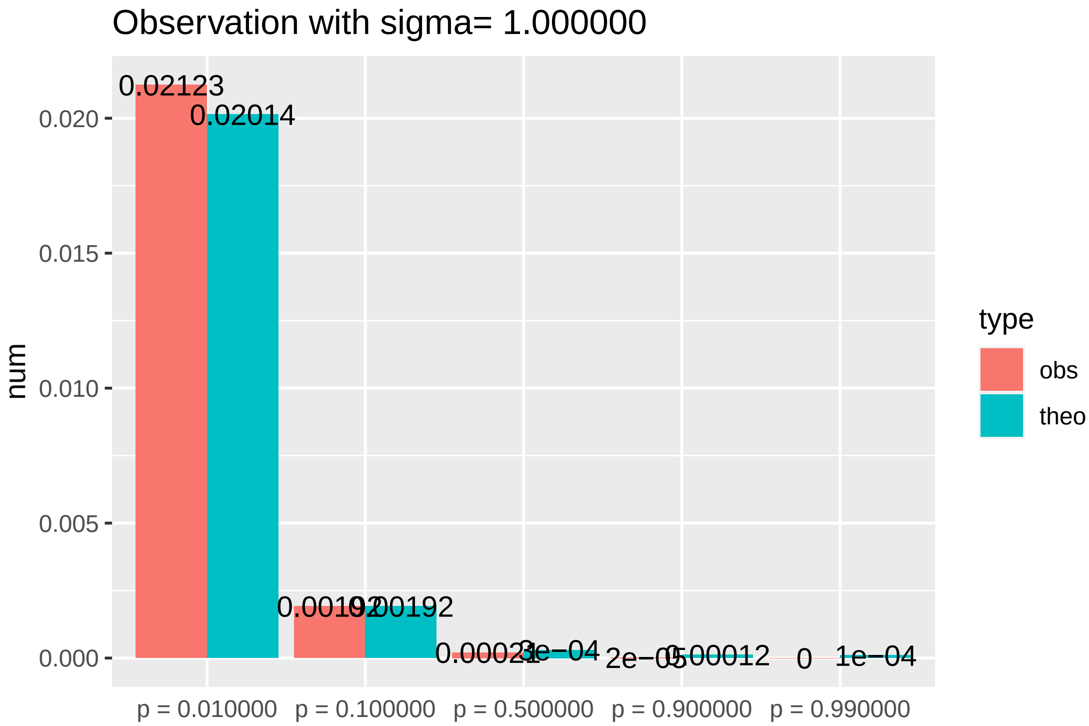


Observation with p= 0.900000
0.00023
0.00012
3e−05
2e−05
1e−05
1e−05
0.00020
0.00015
0.00010
0.00005
0.00000
num
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo

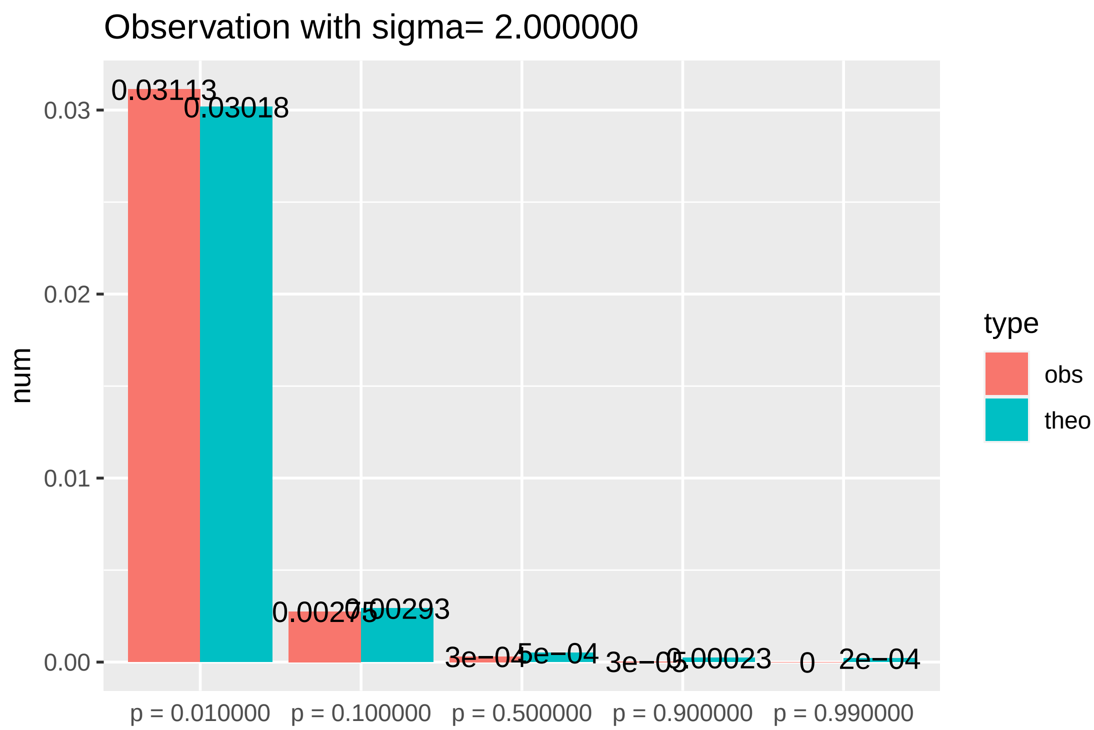



```r
for (i in (c(1:length(sigma_epi2)))){
a<-ggplot(result1 %>% filter(sigma==sprintf('sigma_epi2 = %f',sigma_epi2[i])))+
    geom_col(aes(prob,num,fill=type),position='dodge',width = 0.9)+
  geom_text(aes(prob,num,label=round(as.numeric(num),5),group=type),
            position = position_dodge(width = 0.9))+
  theme(axis.title.x=element_blank(),
        axis.ticks.x=element_blank(),
        plot.margin = unit(c(1,1,1,1), "cm"))+
  ggtitle(sprintf('Observation with sigma= %f',sigma_epi2[i]))
plot(a)
}
```


Observation with sigma= 0.000000
num
0.0100
0.0075
0.0050
0.0025
0.0000
0.01055
0.01011
9e−04
0.00092
1e−04
1e−04
1e−05
1e−05
0
0
p = 0.010000
p = 0.100000
p = 0.500000
p = 0.900000
p = 0.990000
type
obs
theo



Observation with sigma= 1.000000
0.02123
0.02014
0.00192
0.00192
0.00021
3e−04
2e−05
0.00012
0
1e−04
0.020
0.015
0.010
0.005
0.000
num
p = 0.010000
p = 0.100000
p = 0.500000
p = 0.900000
p = 0.990000
type
obs
theo



prob	sigma	obs	theo
p = 0.010000	sigma_epi2 = 0.000000	0.0105500840337253	0.0101093961864815
p = 0.100000	sigma_epi2 = 0.000000	0.000901481187646336	0.000919036016952866
p = 0.500000	sigma_epi2 = 0.000000	0.000102662362139251	0.000102115112994763
p = 0.900000	sigma_epi2 = 0.000000	1.1466816865127e-05	1.13461236660848e-05
p = 0.990000	sigma_epi2 = 0.000000	9.92683646015324e-07	1.03146578782589e-06
p = 0.010000	sigma_epi2 = 1.000000	0.0212336539129811	0.0201427301533014
p = 0.100000	sigma_epi2 = 1.000000	0.00192197993445238	0.00192236941363486
p = 0.500000	sigma_epi2 = 1.000000	0.00021171437794942	0.000302781792331161
p = 0.900000	sigma_epi2 = 1.000000	2.31220067686445e-05	0.000122827612186306
p = 0.990000	sigma_epi2 = 1.000000	2.12221880482516e-06	0.000102378273533482
p = 0.010000	sigma_epi2 = 2.000000	0.0311268197717426	0.0301760641201214
p = 0.100000	sigma_epi2 = 2.000000	0.00275301807570909	0.00292570281031685
p = 0.500000	sigma_epi2 = 2.000000	0.000302215103249437	0.000503448471667559
p = 0.900000	sigma_epi2 = 2.000000	3.43376580837948e-05	0.000234309100706527
p = 0.990000	sigma_epi2 = 2.000000	3.02338995876932e-06	0.000203725081279137

prob	sigma	time
p = 0.010000	sigma_epi2 = 0.000000	3.84671401977539
p = 0.100000	sigma_epi2 = 0.000000	4.10406088829041
p = 0.500000	sigma_epi2 = 0.000000	4.93380403518677
p = 0.900000	sigma_epi2 = 0.000000	4.83507204055786
p = 0.990000	sigma_epi2 = 0.000000	4.82805800437927
p = 0.010000	sigma_epi2 = 1.000000	3.88757801055908
p = 0.100000	sigma_epi2 = 1.000000	4.23616790771484
p = 0.500000	sigma_epi2 = 1.000000	4.98786616325378
p = 0.900000	sigma_epi2 = 1.000000	4.85099601745605
p = 0.990000	sigma_epi2 = 1.000000	4.81013679504395
p = 0.010000	sigma_epi2 = 2.000000	3.90356206893921
p = 0.100000	sigma_epi2 = 2.000000	4.0661609172821
p = 0.500000	sigma_epi2 = 2.000000	5.04550719261169
p = 0.900000	sigma_epi2 = 2.000000	4.82110381126404
p = 0.990000	sigma_epi2 = 2.000000	4.83707118034363
time in total	time in total	1.16278588374456