Progress Report

Victor Huang

November 17, 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.
```

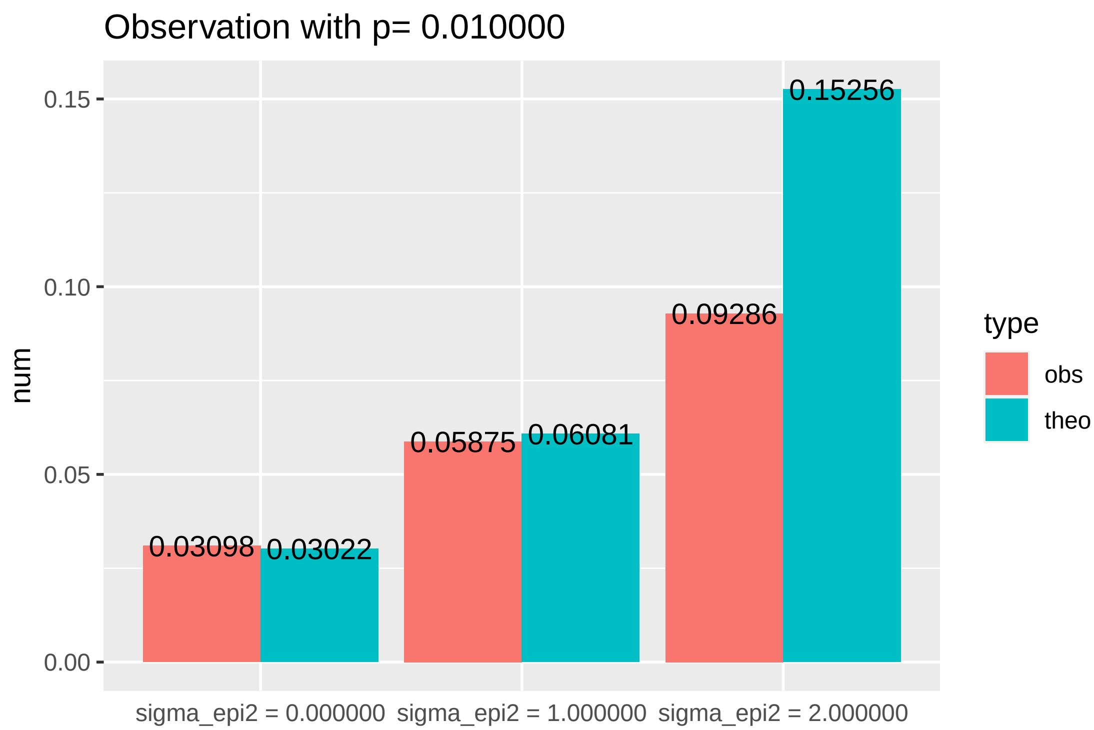
```r
# setting parameters
sigma_e2<-1
N<-10000
beta<-2
nSims<-10000

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,'uniform')
ei<-pop_gen(e_para,N,'normal')

yi<-xi*beta+ei



result<-NULL

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

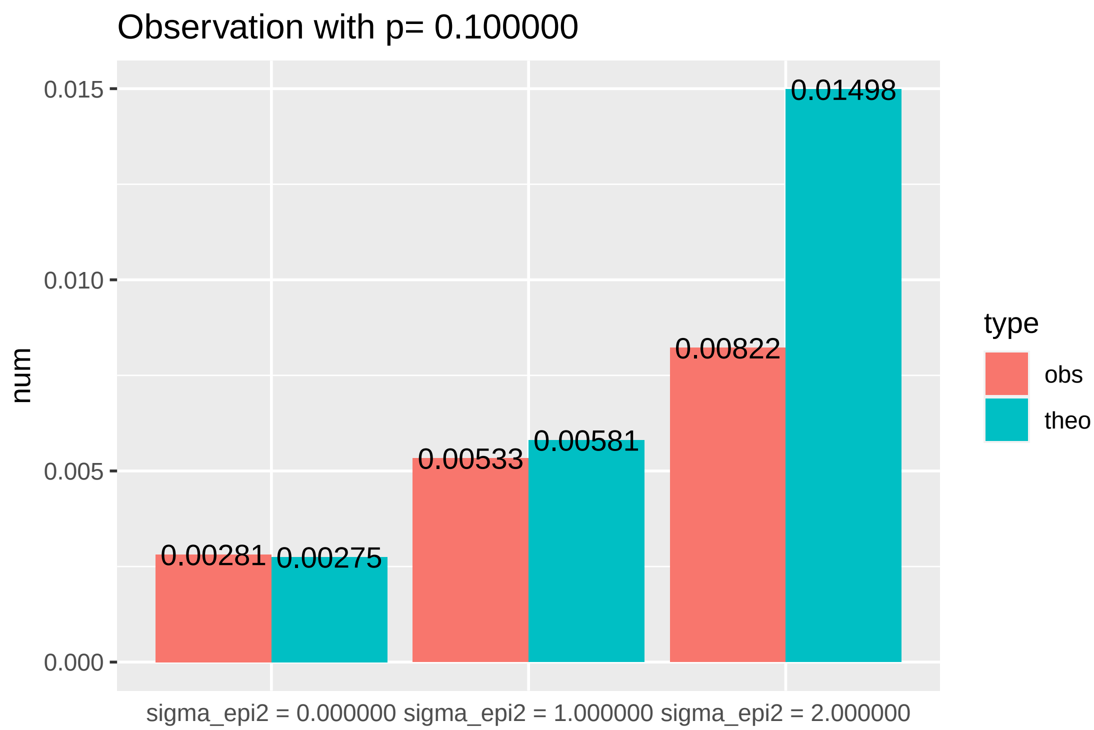
```r
# independent error epislon.

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]**2/(sum(xi**2))+((1-p[i])/p[i])*
      ((sum(xi**2*ei**2))/((sum(xi**2))**2)+sigma_epi2[j]**2/(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)
  }
}
```

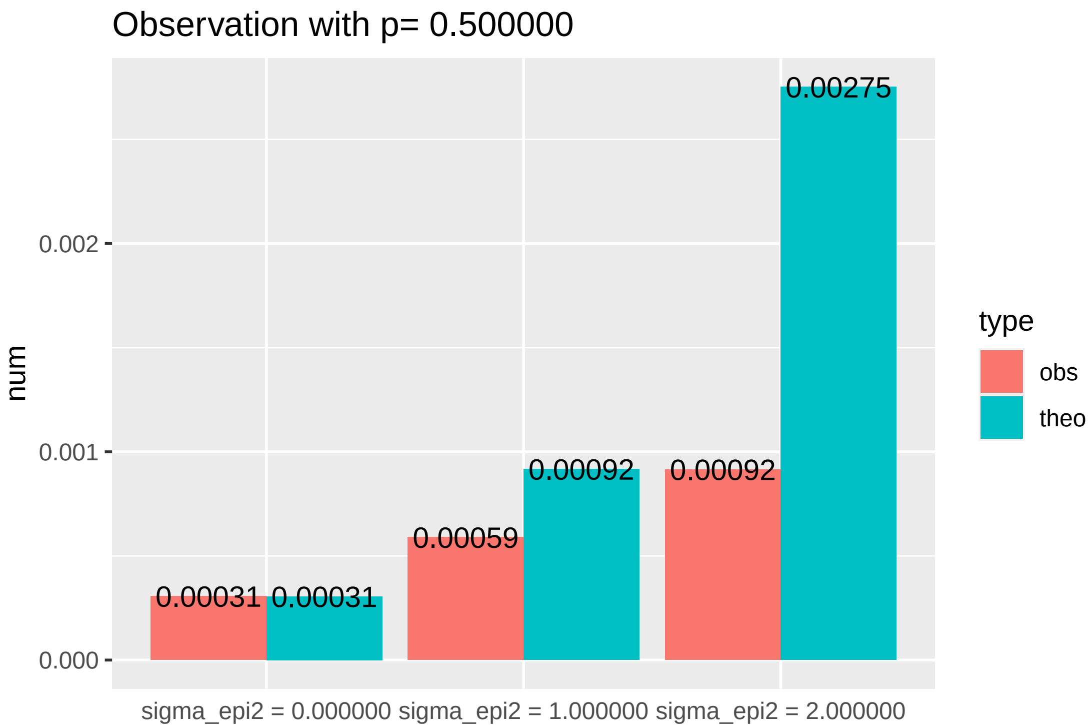
```r
result1<-as.data.frame(result) %>% gather(type,num,obs:theo) %>% type_convert()
```

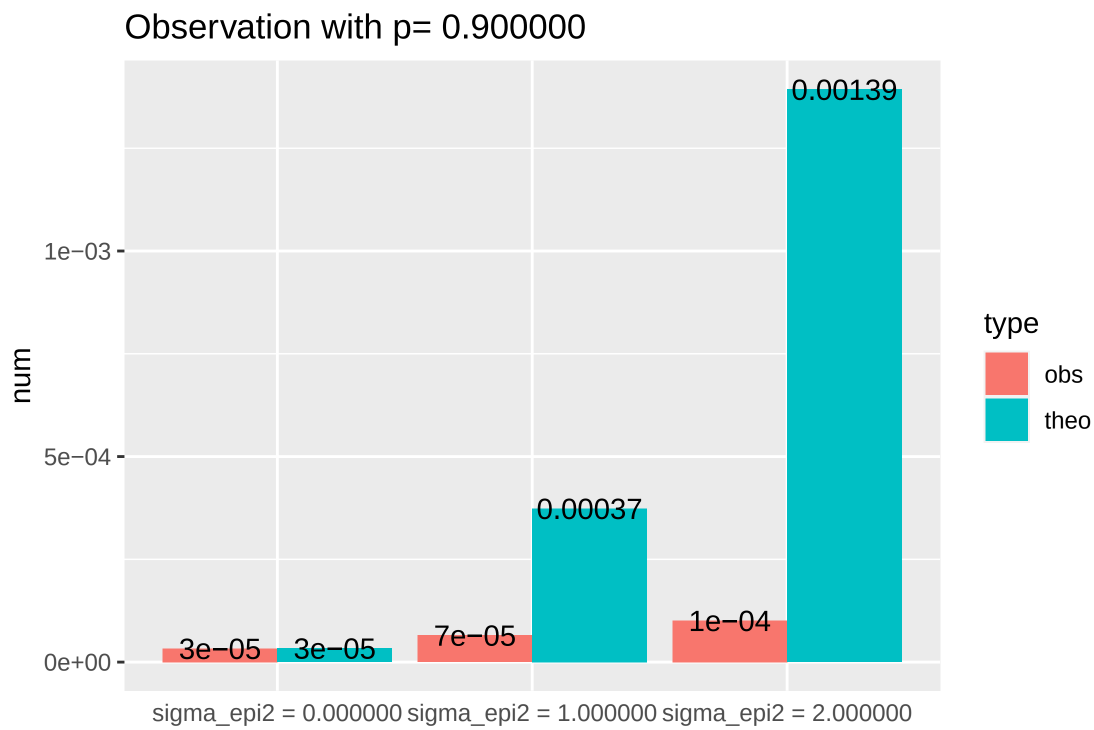
```
## Parsed with column specification:
## cols(
##   prob = col_character(),
##   sigma = col_character(),
##   time = col_double(),
##   type = col_character(),
##   num = col_double()
## )
```

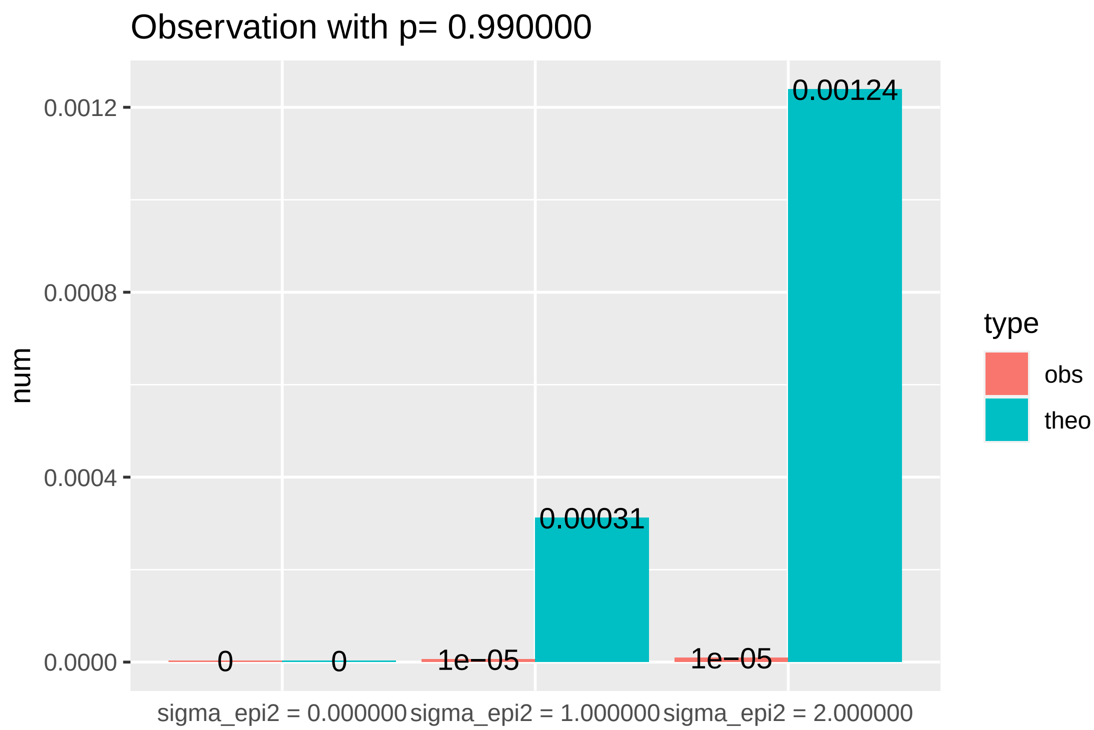
```r
for (i in (c(1:length(p)))){
a<-ggplot(result1 %>% filter(prob==sprintf('p = %f',p[i])))+
    geom_col(aes(sigma,num,fill=type),position='dodge',width = 0.9)+
  geom_text(aes(sigma,num,label=round(as.numeric(num),5),group=type),
            position = position_dodge(width = 0.9))+
  theme(axis.title.x=element_blank(),
        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
0.15
0.10
0.05
0.00
0.03098
0.03022
0.05875
0.06081
0.09286
0.15256
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo

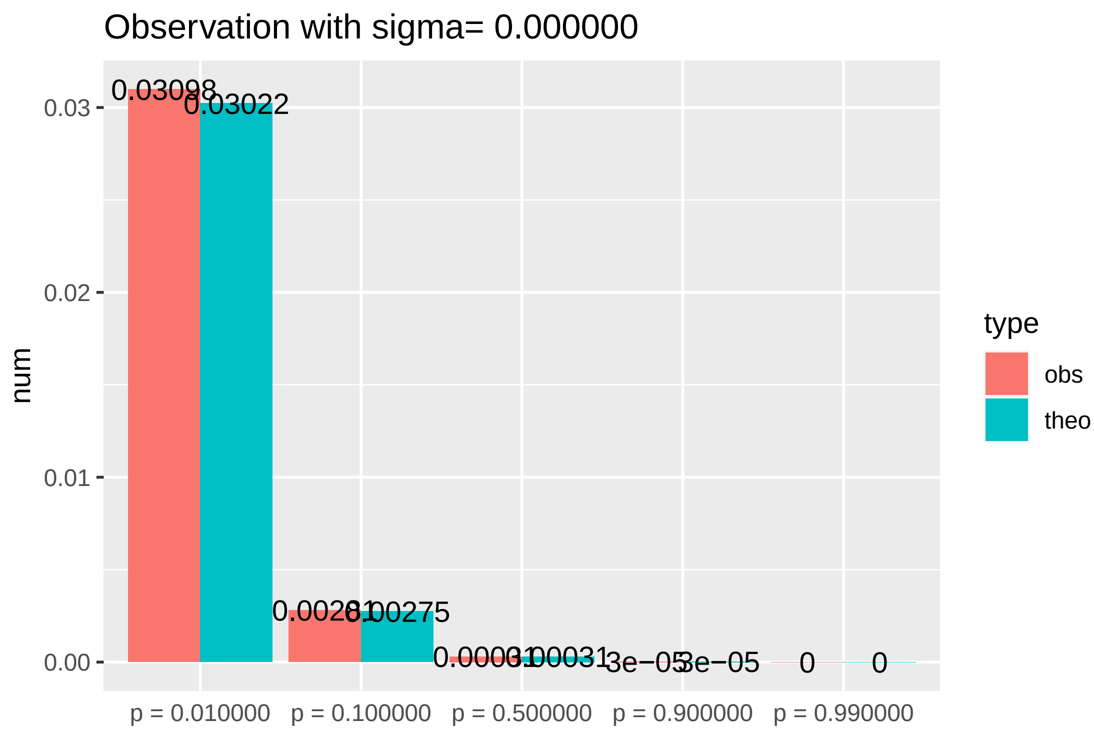

Observation with p= 0.100000
num
0.015
0.010
0.005
0.000
0.00281
0.00275
0.00533
0.00581
0.00822
0.01498
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo

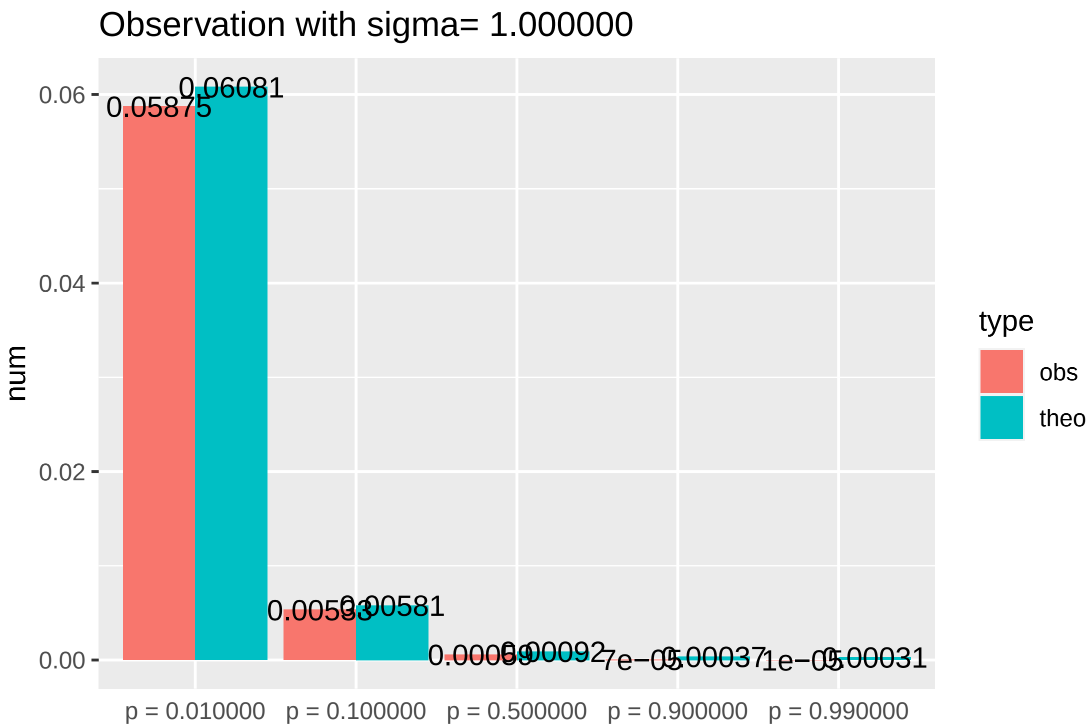

Observation with p= 0.500000
0.00275
0.002
num
0.001
0.00092
0.00092
0.00059
0.00031
0.00031
0.000
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo

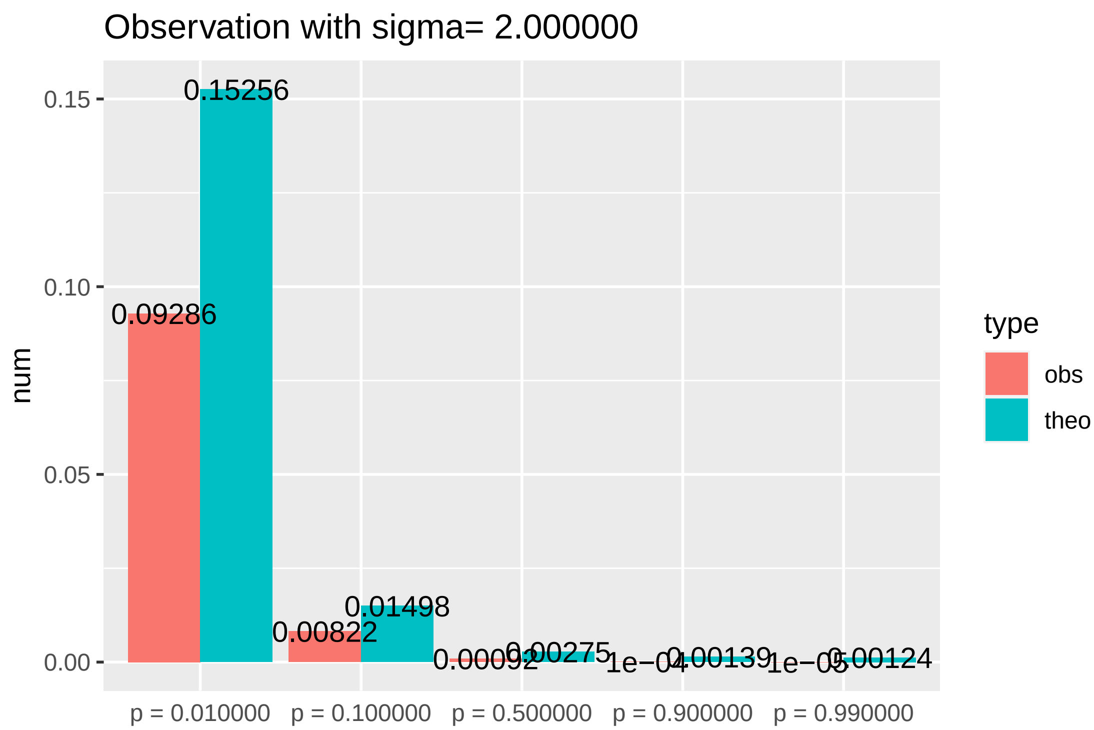

Observation with p= 0.900000
num
1e−03
5e−04
0e+00
0.00139
0.00037
3e−05
3e−05
7e−05
1e−04
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo




```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
0.03098
0.03022
0.00281
0.00275
0.00031
0.00031
3e−05
3e−05
0
0
0.03
0.02
0.01
0.00
num
p = 0.010000
p = 0.100000
p = 0.500000
p = 0.900000
p = 0.990000
type
obs
theo


Observation with sigma= 1.000000
num
0.06
0.04
0.02
0.00
0.05875
0.06081
0.00533
0.00581
0.00059
0.00092
7e−05
0.00037
1e−05
0.00031
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	0.0309820929830577	0.0302233936639482
p = 0.100000	sigma_epi2 = 0.000000	0.0028126120685581	0.00274758124217711
p = 0.500000	sigma_epi2 = 0.000000	0.000306686605479657	0.000305286804686345
p = 0.900000	sigma_epi2 = 0.000000	3.3258308342442e-05	3.39207560762606e-05
p = 0.990000	sigma_epi2 = 0.000000	3.16706115708978e-06	3.08370509784188e-06
p = 0.010000	sigma_epi2 = 1.000000	0.0587485262359455	0.0608084513836468
p = 0.100000	sigma_epi2 = 1.000000	0.00533239649855778	0.00580608701414697
p = 0.500000	sigma_epi2 = 1.000000	0.000589627313652621	0.000916987959080318
p = 0.900000	sigma_epi2 = 1.000000	6.5422928552495e-05	0.000373754730739579
p = 0.990000	sigma_epi2 = 1.000000	5.99449038981721e-06	0.000312023682064495
p = 0.010000	sigma_epi2 = 2.000000	0.0928646403774771	0.152563624542743
p = 0.100000	sigma_epi2 = 2.000000	0.00821917283742358	0.0149816043300566
p = 0.500000	sigma_epi2 = 2.000000	0.00091532604629587	0.00275209142226224
p = 0.900000	sigma_epi2 = 2.000000	0.000101174283052235	0.00139325665472953
p = 0.990000	sigma_epi2 = 2.000000	9.23704425158607e-06	0.00123884361296445

prob	sigma	time
p = 0.010000	sigma_epi2 = 0.000000	11.5573019981384
p = 0.100000	sigma_epi2 = 0.000000	12.1345460414886
p = 0.500000	sigma_epi2 = 0.000000	16.3247919082642
p = 0.900000	sigma_epi2 = 0.000000	20.2859809398651
p = 0.990000	sigma_epi2 = 0.000000	21.3644959926605
p = 0.010000	sigma_epi2 = 1.000000	11.1322000026703
p = 0.100000	sigma_epi2 = 1.000000	12.0866811275482
p = 0.500000	sigma_epi2 = 1.000000	16.1598169803619
p = 0.900000	sigma_epi2 = 1.000000	20.1261849403381
p = 0.990000	sigma_epi2 = 1.000000	21.1203570365906
p = 0.010000	sigma_epi2 = 2.000000	11.1382100582123
p = 0.100000	sigma_epi2 = 2.000000	12.0130701065063
p = 0.500000	sigma_epi2 = 2.000000	16.2375848293304
p = 0.900000	sigma_epi2 = 2.000000	20.1389660835266
p = 0.990000	sigma_epi2 = 2.000000	21.3441178798676
time in total	time in total	4.12616151571274