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.
```

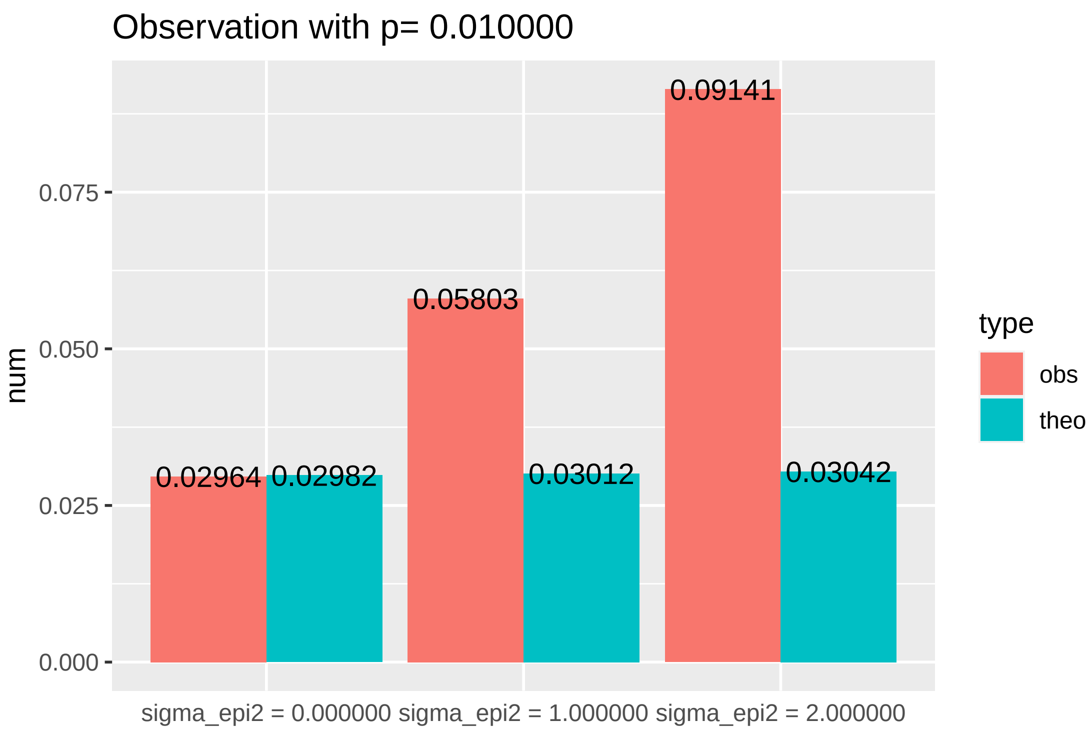
```r
# setting parameters
sigma_e2<-1
N<-10000
beta<-2
nSims<-10000

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
```

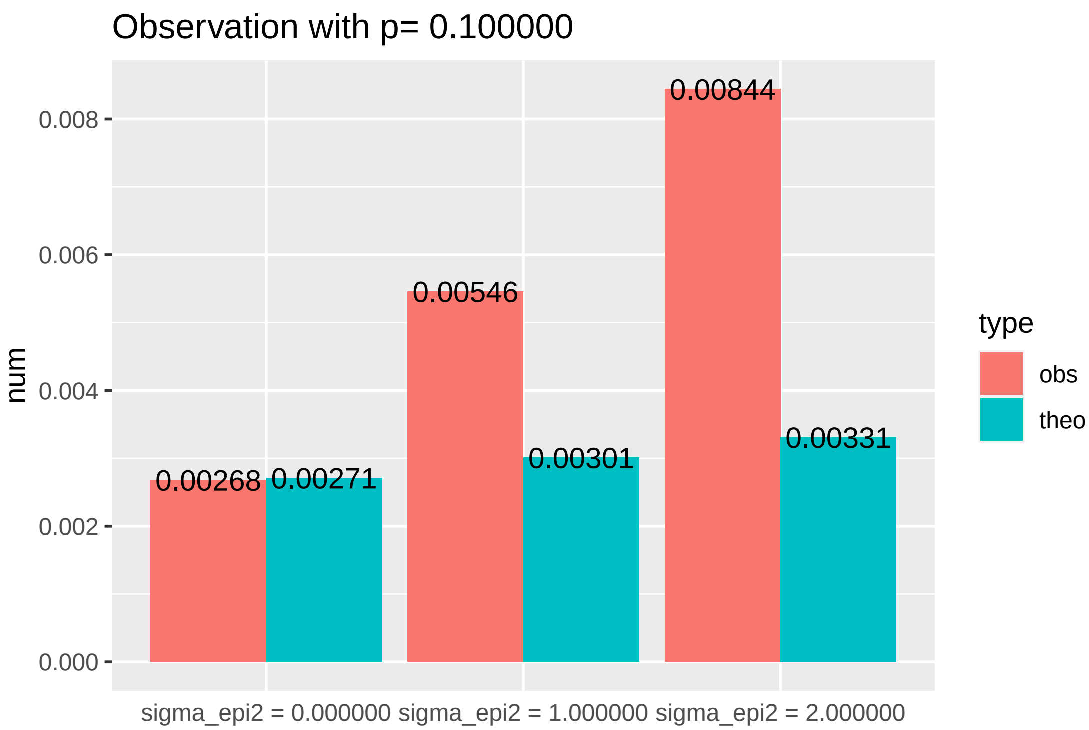
```r
# independent error epislon.

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

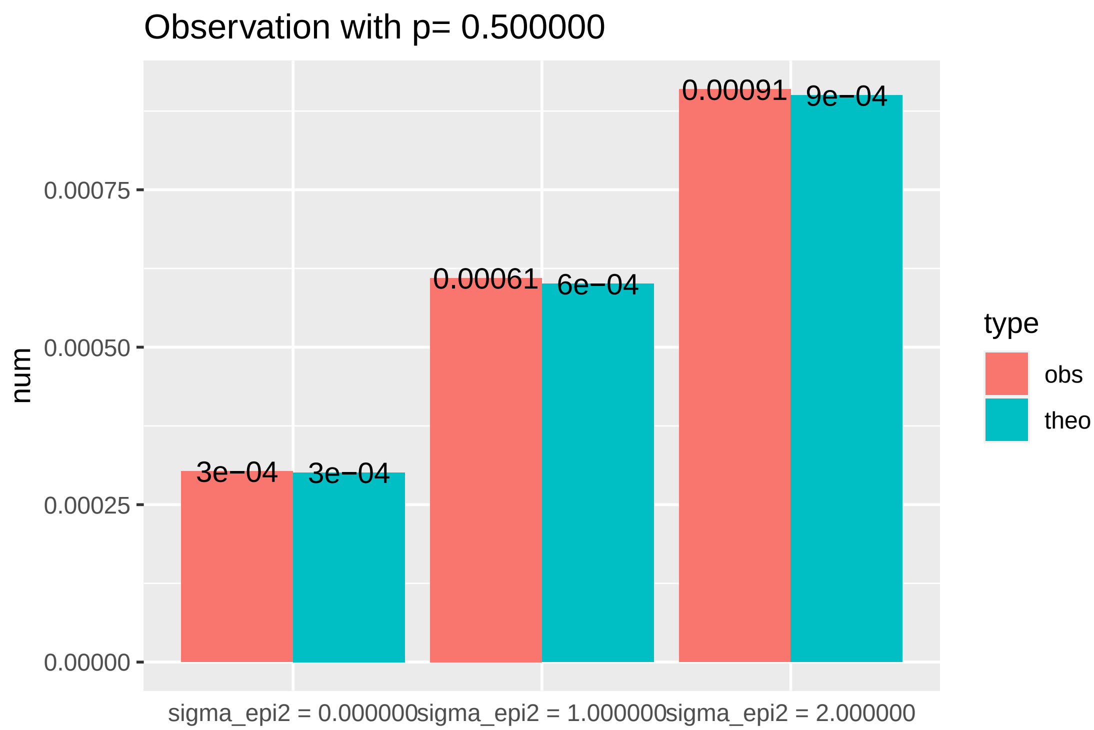
```r
result1<-as.data.frame(result) %>% gather(type,num,obs:theo) %>% type_convert()
```

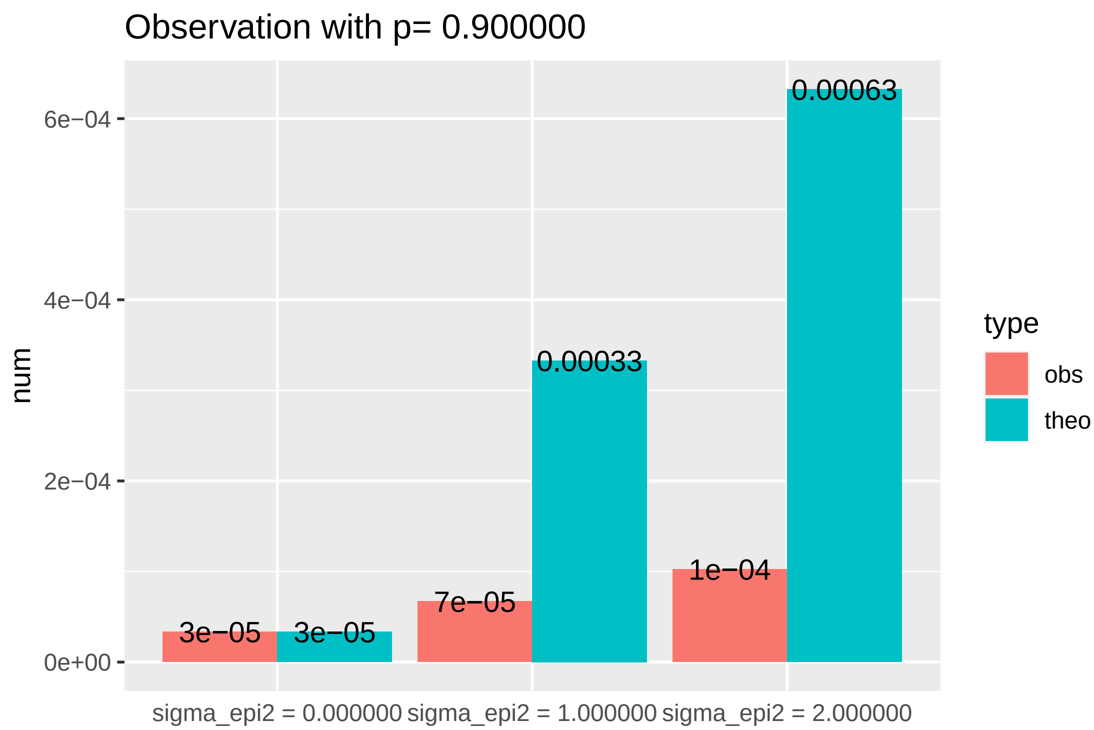
```
## Parsed with column specification:
## cols(
##   prob = col_character(),
##   sigma = col_character(),
##   time = col_double(),
##   type = col_character(),
##   num = col_double()
## )
```

```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)
}
```

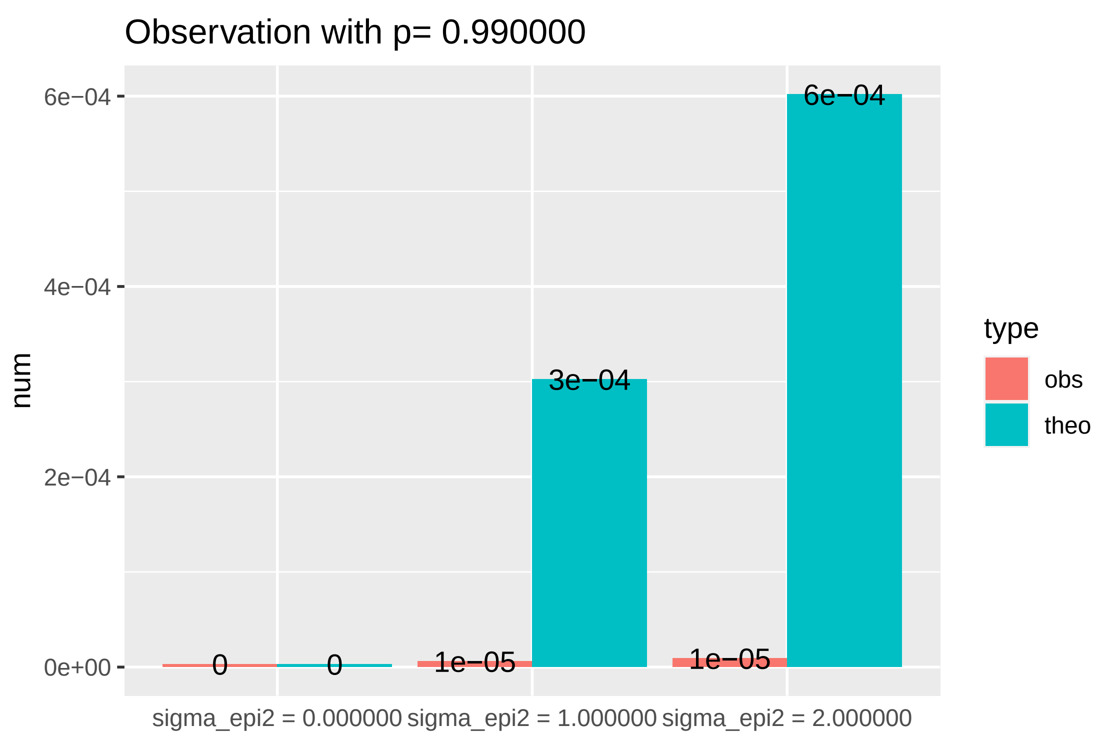

Observation with p= 0.010000
num
0.000
0.025
0.050
0.075
0.02964
0.02982
0.05803
0.03012
0.09141
0.03042
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo


Observation with p= 0.100000
num
0.008
0.006
0.004
0.002
0.000
0.00844
0.00546
0.00268
0.00271
0.00301
0.00331
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo

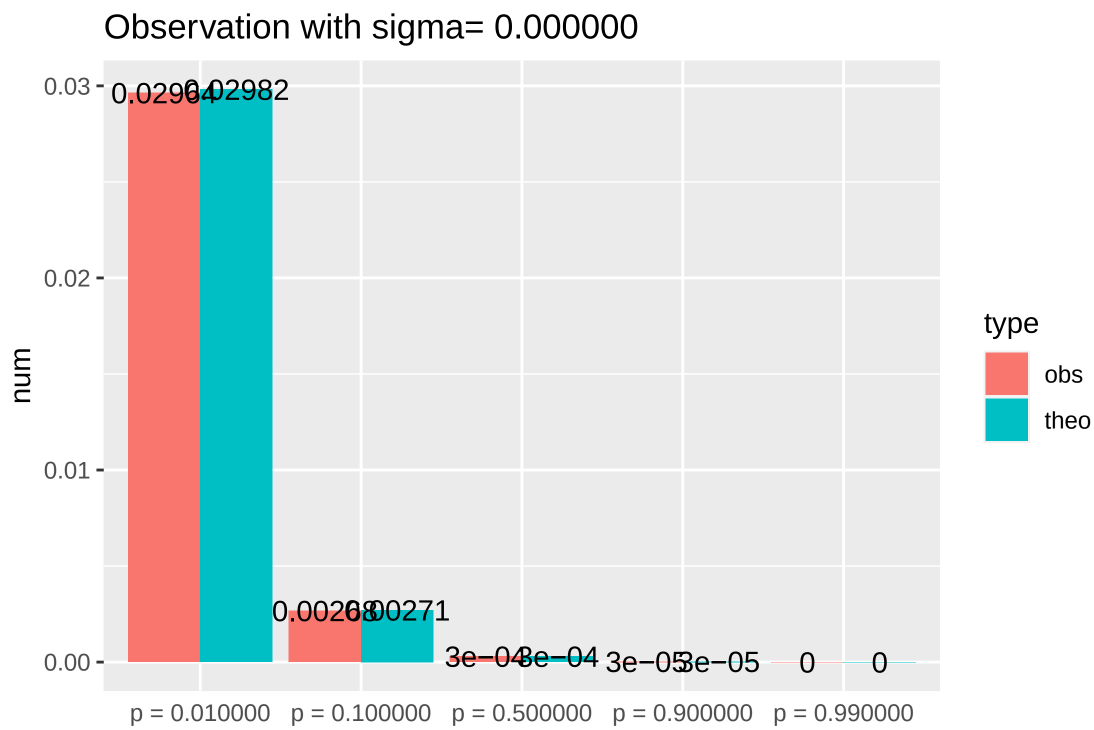

Observation with p= 0.500000
num
0.00075
0.00050
0.00025
0.00000
3e−04
3e−04
0.00061
6e−04
0.00091
9e−04
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo



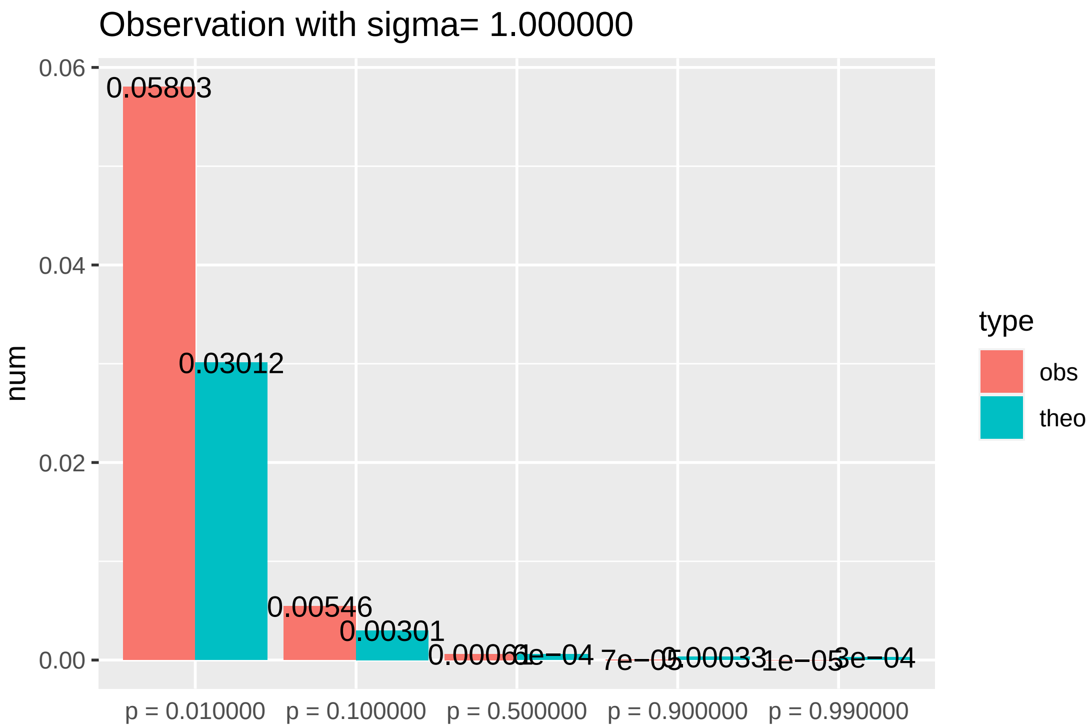
Observation with p= 0.900000
num
6e−04
4e−04
2e−04
0e+00
3e−05
3e−05
7e−05
0.00033
1e−04
0.00063
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo

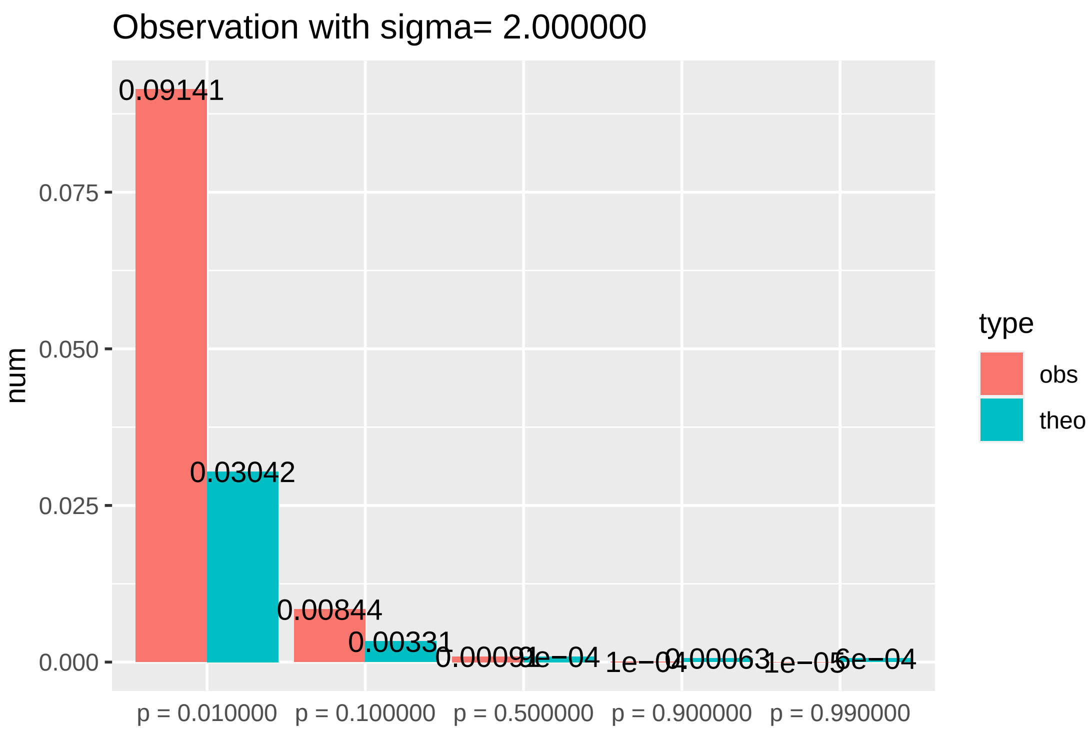



```
for (i in (c(1:length(sigma_epi2)))){
a<-ggplot(result1 %>% filter(sigma==sprintf('sigma_epi2 = %f',sigma_epi2[i])))+
    geom_col(aes(prob,num,fill=type),position='dodge',width = 0.9)+
  geom_text(aes(prob,num,label=round(as.numeric(num),5),group=type),
            position = position_dodge(width = 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
0.02964
0.02982
0.00268
0.00271
3e−04
3e−04
3e−05
3e−05
0
0
num
0.03
0.02
0.01
0.00
p = 0.010000
p = 0.100000
p = 0.500000
p = 0.900000
p = 0.990000
type
obs
theo


Observation with sigma= 1.000000
0.06
0.04
0.02
0.00
num
0.05803
0.03012
0.00546
0.00301
0.00061
1e−04
7e−05
0.00033
1e−05
3e−04
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	0.0296398600749798	0.029818226013555
p = 0.100000	sigma_epi2 = 0.000000	0.00267956093873299	0.0027107478194141
p = 0.500000	sigma_epi2 = 0.000000	0.0003028899894226	0.000301194202157122
p = 0.900000	sigma_epi2 = 0.000000	3.35704889090169e-05	3.34660224619024e-05
p = 0.990000	sigma_epi2 = 0.000000	3.14393782217354e-06	3.04236567835477e-06
p = 0.010000	sigma_epi2 = 1.000000	0.0580273295277495	0.0301176602174658
p = 0.100000	sigma_epi2 = 1.000000	0.00545862226973184	0.00301018202332482
p = 0.500000	sigma_epi2 = 1.000000	0.00060984009264221	0.000600628406067849
p = 0.900000	sigma_epi2 = 1.000000	6.70799974305426e-05	0.00033290022637263
p = 0.990000	sigma_epi2 = 1.000000	6.04860168772048e-06	0.000302476569589082
p = 0.010000	sigma_epi2 = 2.000000	0.0914136242509702	0.0304170944213765
p = 0.100000	sigma_epi2 = 2.000000	0.00844070521524059	0.00330961622723555
p = 0.500000	sigma_epi2 = 2.000000	0.000909414464277947	0.000900062609978577
p = 0.900000	sigma_epi2 = 2.000000	0.000102568101660501	0.000632334430283357
p = 0.990000	sigma_epi2 = 2.000000	9.20542776949471e-06	0.00060191077349981

prob	sigma	time
p = 0.010000	sigma_epi2 = 0.000000	12.8805449008942
p = 0.100000	sigma_epi2 = 0.000000	13.6524529457092
p = 0.500000	sigma_epi2 = 0.000000	18.4526400566101
p = 0.900000	sigma_epi2 = 0.000000	29.4681849479675
p = 0.990000	sigma_epi2 = 0.000000	25.671352148056
p = 0.010000	sigma_epi2 = 1.000000	13.921758890152
p = 0.100000	sigma_epi2 = 1.000000	15.1973180770874
p = 0.500000	sigma_epi2 = 1.000000	21.329941034317
p = 0.900000	sigma_epi2 = 1.000000	22.3698928356171
p = 0.990000	sigma_epi2 = 1.000000	24.6600370407104
p = 0.010000	sigma_epi2 = 2.000000	12.9942419528961
p = 0.100000	sigma_epi2 = 2.000000	13.2555420398712
p = 0.500000	sigma_epi2 = 2.000000	17.8093581199646
p = 0.900000	sigma_epi2 = 2.000000	20.9838688373566
p = 0.990000	sigma_epi2 = 2.000000	22.9446251392365
time in total	time in total	4.79798349936803