Progress Report

Victor Huang

November 29, 2020

Overview:

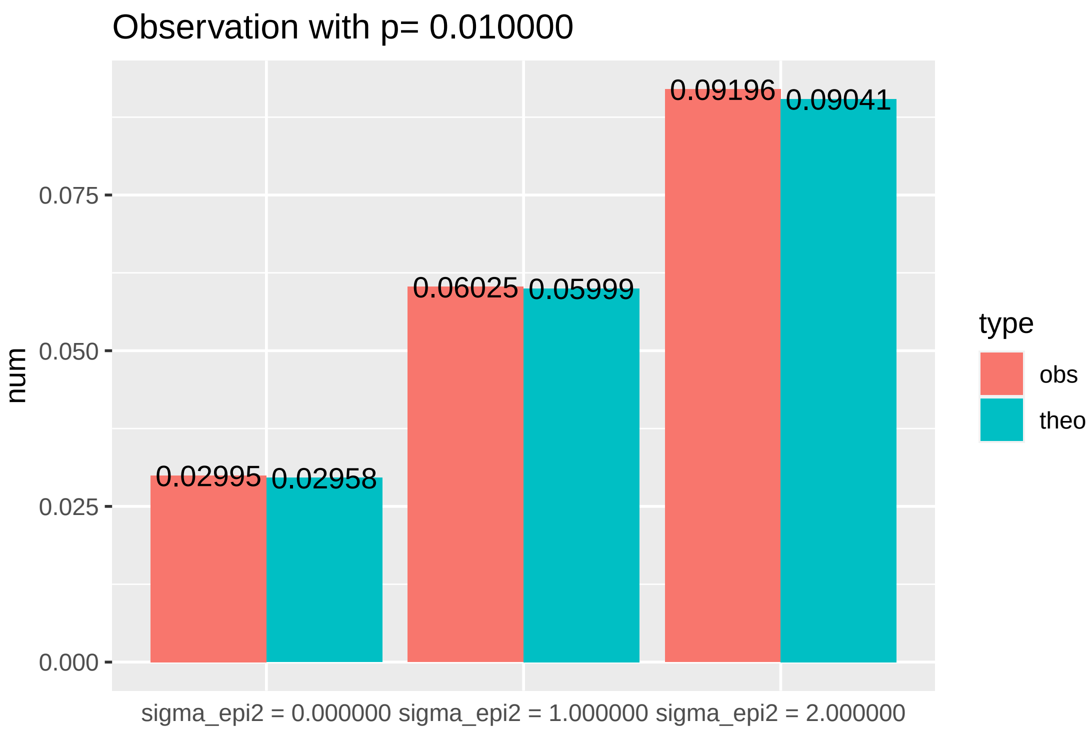
Concerned variables and underlying Monte-Carlo model:

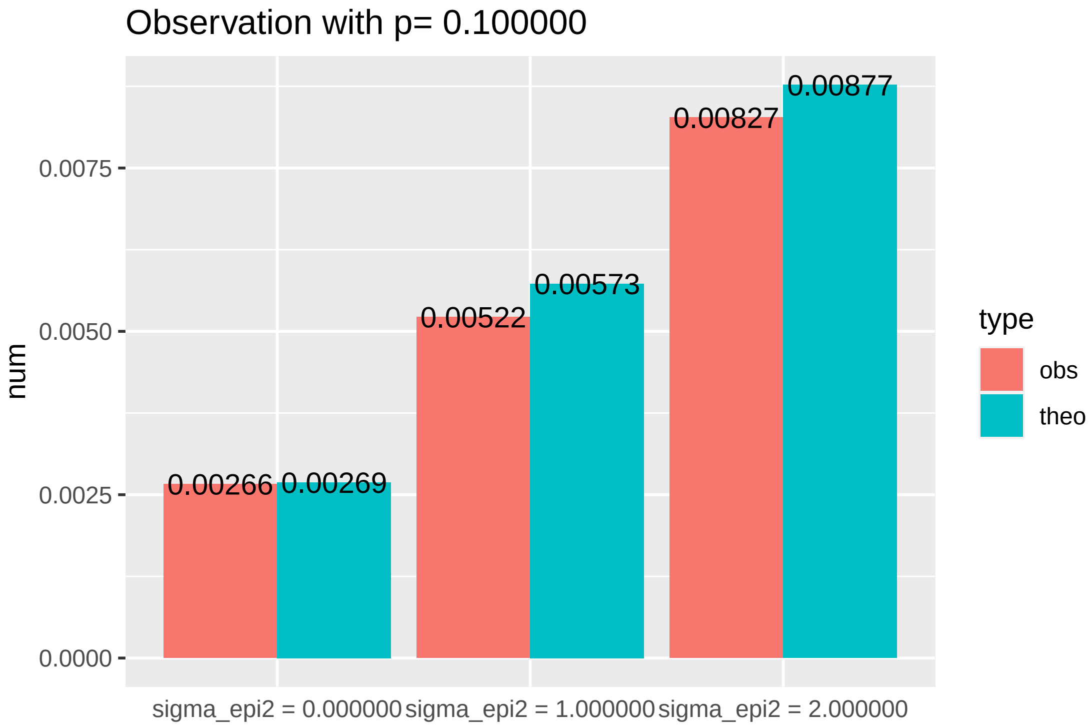
```
nSims: Number of simulations to run for each groups of parameter.
N: Size of population.
p: Probablity an observation is sampled.
sigma_e2: Variance of Model error e.
sigma_epi2: variance of unpredicted error epislon.
yi: Dependent variable without unpredicted error.
xi: Independent variable.
cap_y: 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,'uniform')
ei<-pop_gen(e_para,N,'normal')

yi<-xi*beta+ei



result<-NULL

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

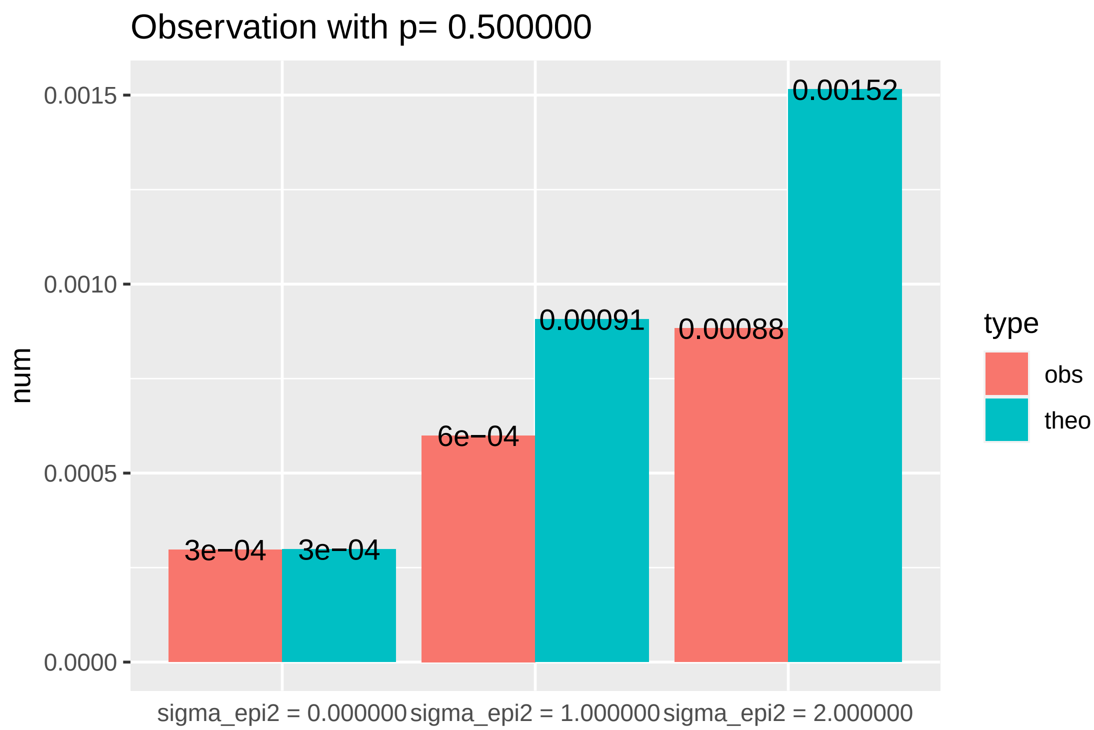
```r
# independent error epislon.

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)+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)
}
```

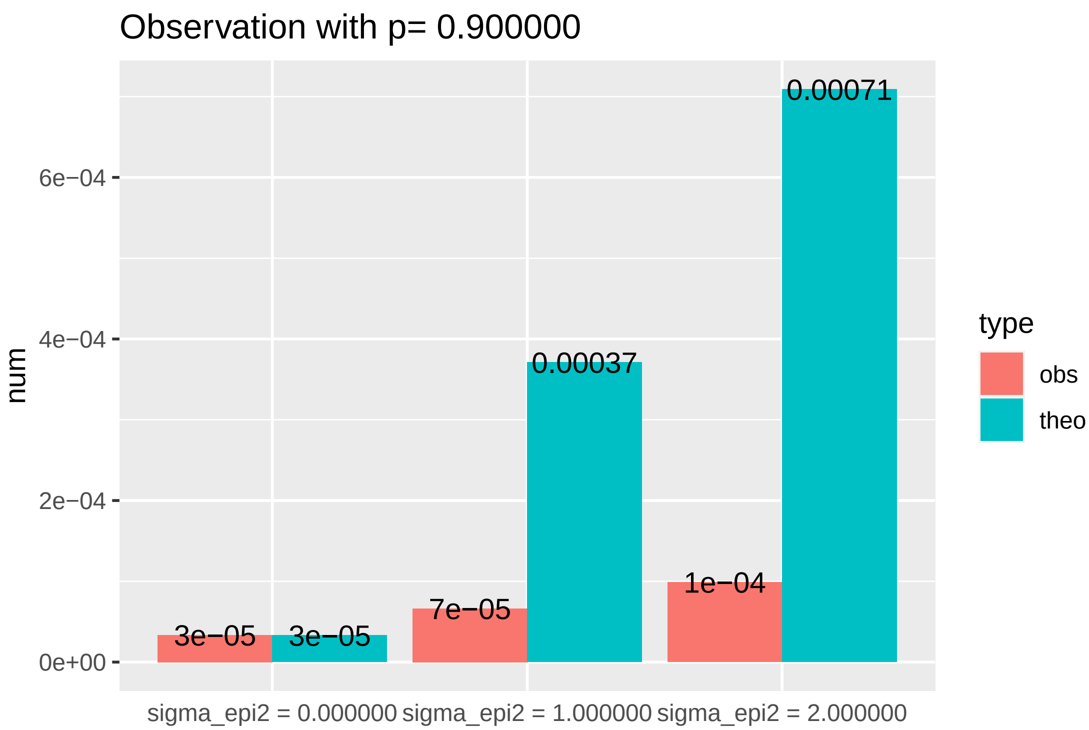

Observation with p= 0.010000
num
0.000
0.025
0.050
0.075
0.02995
0.02958
0.06025
0.05999
0.09196
0.09041
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo

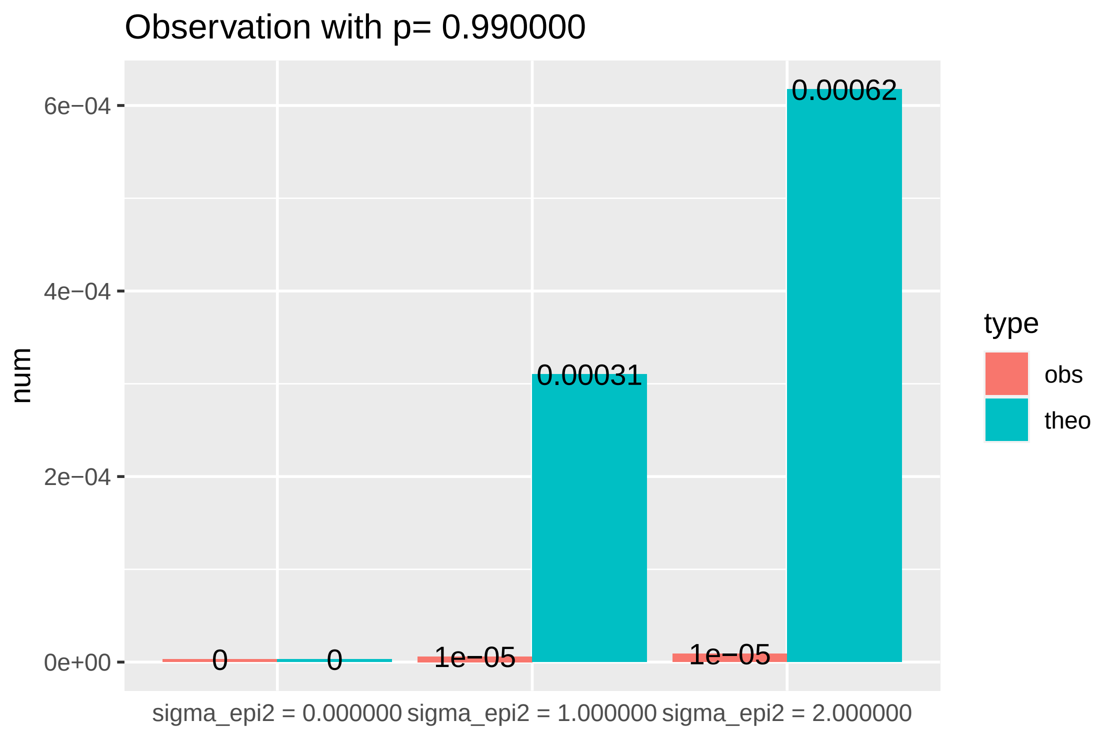


Observation with p= 0.100000
num
0.0000
0.0025
0.0050
0.0075
0.00266
0.00269
0.00522
0.00573
0.00827
0.00877
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo

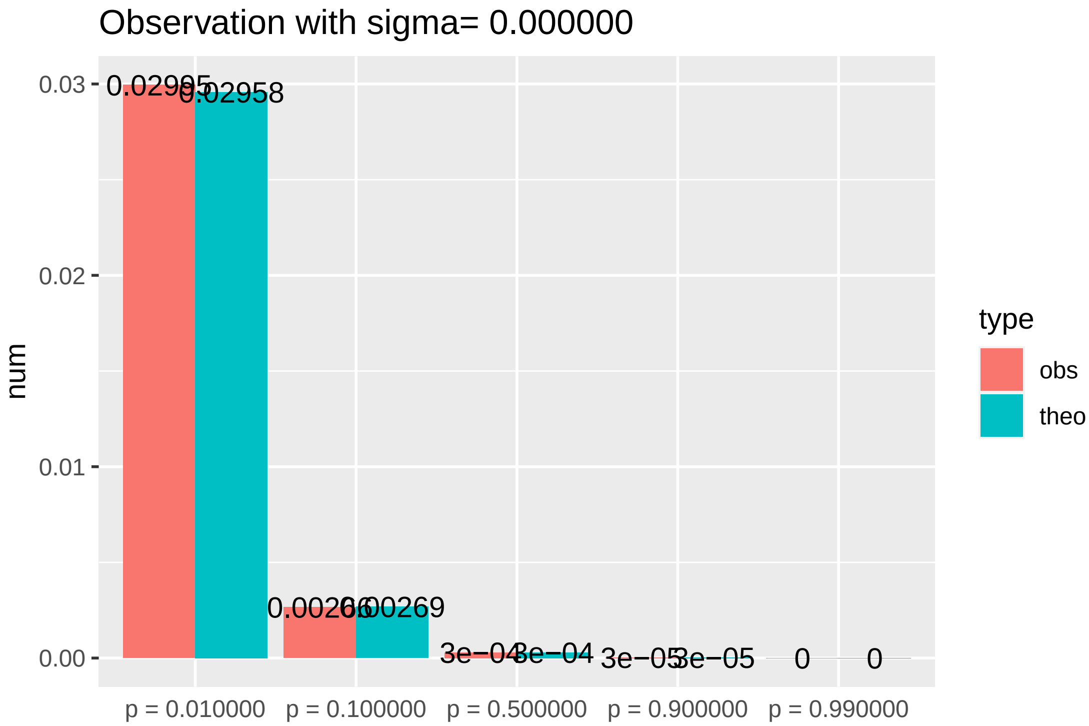


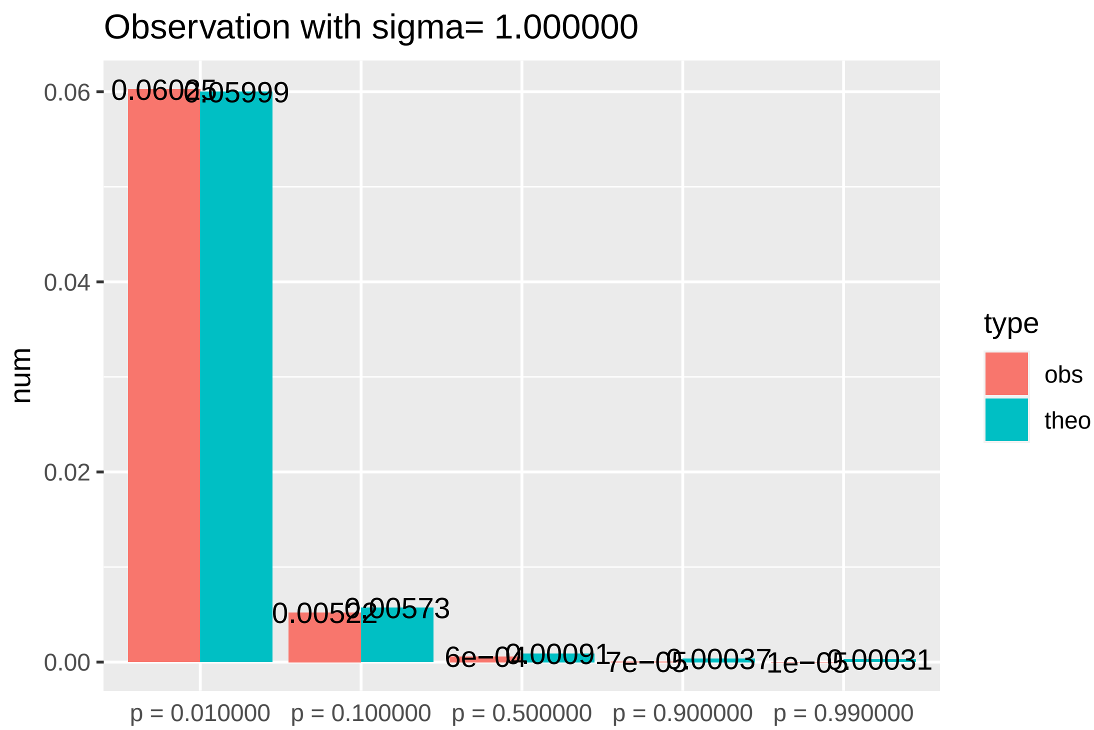
Observation with p= 0.500000
0.0015
0.0010
0.0005
0.0000
num
0.00152
0.00091
0.00088
6e−04
3e−04
3e−04
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo

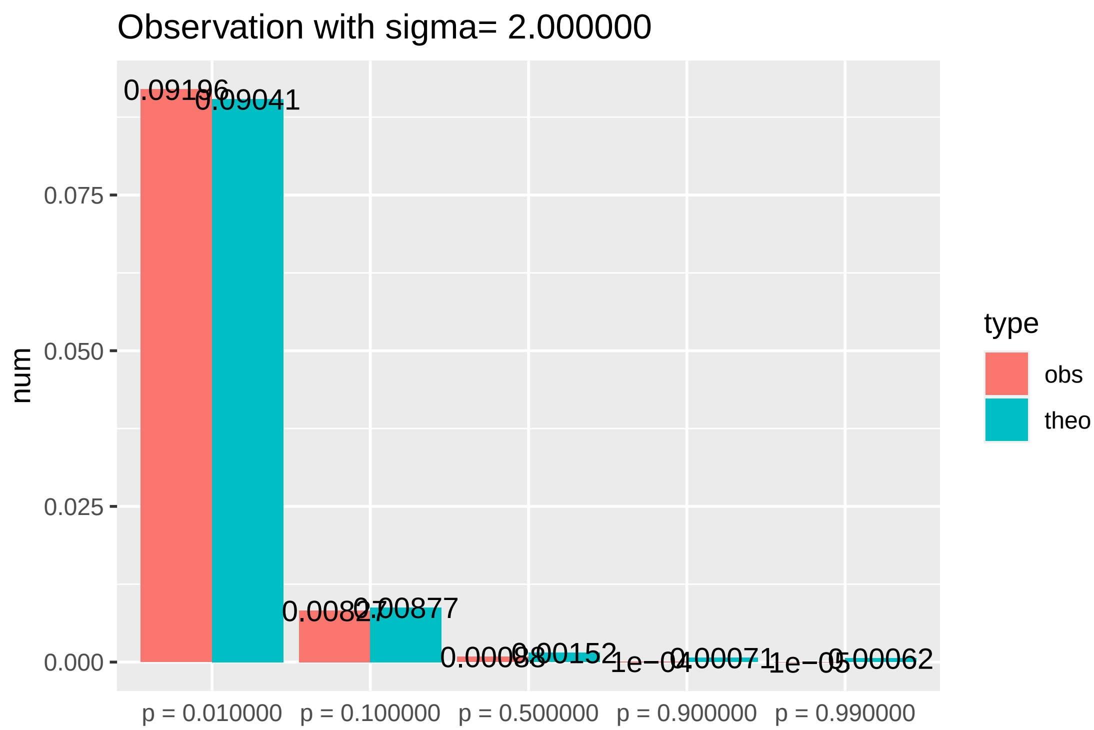

Observation with p= 0.900000
num
6e−04
4e−04
2e−04
0e+00
3e−05
3e−05
7e−05
0.00037
1e−04
0.00071
sigma_epi2 = 0.000000
sigma_epi2 = 1.000000
sigma_epi2 = 2.000000
type
obs
theo




```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
0.03
0.02
0.01
0.00
0.02985
0.02958
0.00266
0.00269
3e−04
3e−04
3e−05
3e−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
num
0.06
0.04
0.02
0.00
0.06025
0.05999
0.00522
0.00573
6e−04
0.00091
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	0.0299514665704794	0.029582626068724
p = 0.100000	sigma_epi2 = 0.000000	0.00266401212574631	0.00268932964261128
p = 0.500000	sigma_epi2 = 0.000000	0.000297435212798823	0.000298814404734586
p = 0.900000	sigma_epi2 = 0.000000	3.35315387378322e-05	3.32016005260651e-05
p = 0.990000	sigma_epi2 = 0.000000	2.96638181366461e-06	3.01832732055138e-06
p = 0.010000	sigma_epi2 = 1.000000	0.0602491900815253	0.0599939639930228
p = 0.100000	sigma_epi2 = 1.000000	0.0052219283226897	0.00573046343504116
p = 0.500000	sigma_epi2 = 1.000000	0.000599667787880193	0.000907041163220562
p = 0.900000	sigma_epi2 = 1.000000	6.62935041135531e-05	0.000371105355240496
p = 0.990000	sigma_epi2 = 1.000000	6.07323212033474e-06	0.000310203558879125
p = 0.010000	sigma_epi2 = 2.000000	0.0919633442408937	0.0904053019173217
p = 0.100000	sigma_epi2 = 2.000000	0.00827404676640942	0.00877159722747104
p = 0.500000	sigma_epi2 = 2.000000	0.000883229388406762	0.00151526792170654
p = 0.900000	sigma_epi2 = 2.000000	9.90431045265223e-05	0.000709009109954928
p = 0.990000	sigma_epi2 = 2.000000	9.01867927477978e-06	0.000617388790437699

prob	sigma	time
p = 0.010000	sigma_epi2 = 0.000000	3.95838212966919
p = 0.100000	sigma_epi2 = 0.000000	4.17284178733826
p = 0.500000	sigma_epi2 = 0.000000	4.94677710533142
p = 0.900000	sigma_epi2 = 0.000000	4.86473202705383
p = 0.990000	sigma_epi2 = 0.000000	4.82209300994873
p = 0.010000	sigma_epi2 = 1.000000	3.77591896057129
p = 0.100000	sigma_epi2 = 1.000000	4.12297701835632
p = 0.500000	sigma_epi2 = 1.000000	4.967689037323
p = 0.900000	sigma_epi2 = 1.000000	4.86766219139099
p = 0.990000	sigma_epi2 = 1.000000	4.87909197807312
p = 0.010000	sigma_epi2 = 2.000000	3.87666702270508
p = 0.100000	sigma_epi2 = 2.000000	4.10000514984131
p = 0.500000	sigma_epi2 = 2.000000	5.02359795570374
p = 0.900000	sigma_epi2 = 2.000000	4.8689820766449
p = 0.990000	sigma_epi2 = 2.000000	4.8064169883728
time in total	time in total	1.16511705319087