Proportion_Studies_Max_0530

Lan Tao

2023-05-30

read data

```
num_greedy_05_m30_pr2_11<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_R
num_patient_05_m30_pr2_11<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_

num_greedy_05_m30_pr2_12<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_R
num_patient_05_m30_pr2_12<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_

num_greedy_05_m30_pr2_13<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_R
num_patient_05_m30_pr2_13<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_

num_greedy_05_m30_pr2_14<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_R
num_patient_05_m30_pr2_14<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_


#-------------------------------------------------------------------------------

num_greedy_05_m50_pr2_11<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_R
num_patient_05_m50_pr2_11<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_

num_greedy_05_m50_pr2_12<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_R
num_patient_05_m50_pr2_12<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_

num_greedy_05_m50_pr2_13<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_R
num_patient_05_m50_pr2_13<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_

num_greedy_05_m50_pr2_14<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_R
num_patient_05_m50_pr2_14<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_


#-------------------------------------------------------------------------------

num_greedy_05_m70_pr2_11<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_R
num_patient_05_m70_pr2_11<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_

num_greedy_05_m70_pr2_12<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_R
num_patient_05_m70_pr2_12<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_

num_greedy_05_m70_pr2_13<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_R
num_patient_05_m70_pr2_13<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_

num_greedy_05_m70_pr2_14<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_R
```

```
num_patient_05_m70_pr2_14<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_

#-------------------------------------------------------------------------------

num_greedy_05_m100_pr2_11<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_
num_patient_05_m100_pr2_11<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation

num_greedy_05_m100_pr2_12<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_
num_patient_05_m100_pr2_12<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation

num_greedy_05_m100_pr2_13<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_
num_patient_05_m100_pr2_13<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation

num_greedy_05_m100_pr2_14<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation_
num_patient_05_m100_pr2_14<-read_excel("~/Desktop/Matching Market/Simulation/Simulation_0510/Simulation
```

```
# write a function to update the dataframe s.t. s=min{n_A/n_B, n_B/n_A}

update_dta<-function(dta){
  new_dta<-dta
  new_dta$prop_2<-new_dta$num_B/new_dta$num_A
  prop_s<-apply(new_dta[,5:6],1,min)
  new_dta$prop_s<-as.double(prop_s)
  new_dta
}
```

update data

```
num_greedy_05_m30_pr2_11<-update_dta(num_greedy_05_m30_pr2_11)
num_patient_05_m30_pr2_11<-update_dta(num_patient_05_m30_pr2_11)

num_greedy_05_m30_pr2_12<-update_dta(num_greedy_05_m30_pr2_12)
num_patient_05_m30_pr2_12<-update_dta(num_patient_05_m30_pr2_12)

num_greedy_05_m30_pr2_13<-update_dta(num_greedy_05_m30_pr2_13)
num_patient_05_m30_pr2_13<-update_dta(num_patient_05_m30_pr2_13)

num_greedy_05_m30_pr2_14<-update_dta(num_greedy_05_m30_pr2_14)
num_patient_05_m30_pr2_14<-update_dta(num_patient_05_m30_pr2_14)


#-------------------------------------------------------------------------------

num_greedy_05_m50_pr2_11<-update_dta(num_greedy_05_m50_pr2_11)
num_patient_05_m50_pr2_11<-update_dta(num_patient_05_m50_pr2_11)

num_greedy_05_m50_pr2_12<-update_dta(num_greedy_05_m50_pr2_12)
num_patient_05_m50_pr2_12<-update_dta(num_patient_05_m50_pr2_12)

num_greedy_05_m50_pr2_13<-update_dta(num_greedy_05_m50_pr2_13)
num_patient_05_m50_pr2_13<-update_dta(num_patient_05_m50_pr2_13)
```

```
num_greedy_05_m50_pr2_14<-update_dta(num_greedy_05_m50_pr2_14)
num_patient_05_m50_pr2_14<-update_dta(num_patient_05_m50_pr2_14)

#-------------------------------------------------------------------------------------

num_greedy_05_m70_pr2_11<-update_dta(num_greedy_05_m70_pr2_11)
num_patient_05_m70_pr2_11<-update_dta(num_patient_05_m70_pr2_11)

num_greedy_05_m70_pr2_12<-update_dta(num_greedy_05_m70_pr2_12)
num_patient_05_m70_pr2_12<-update_dta(num_patient_05_m70_pr2_12)

num_greedy_05_m70_pr2_13<-update_dta(num_greedy_05_m70_pr2_13)
num_patient_05_m70_pr2_13<-update_dta(num_patient_05_m70_pr2_13)

num_greedy_05_m70_pr2_14<-update_dta(num_greedy_05_m70_pr2_14)
num_patient_05_m70_pr2_14<-update_dta(num_patient_05_m70_pr2_14)

#-------------------------------------------------------------------------------------

num_greedy_05_m100_pr2_11<-update_dta(num_greedy_05_m100_pr2_11)
num_patient_05_m100_pr2_11<-update_dta(num_patient_05_m100_pr2_11)

num_greedy_05_m100_pr2_12<-update_dta(num_greedy_05_m100_pr2_12)
num_patient_05_m100_pr2_12<-update_dta(num_patient_05_m100_pr2_12)

num_greedy_05_m100_pr2_13<-update_dta(num_greedy_05_m100_pr2_13)
num_patient_05_m100_pr2_13<-update_dta(num_patient_05_m100_pr2_13)

num_greedy_05_m100_pr2_14<-update_dta(num_greedy_05_m100_pr2_14)
num_patient_05_m100_pr2_14<-update_dta(num_patient_05_m100_pr2_14)
```

Part 1 - plot the distribution of s

2.1 Greedy algorithm plot

```
p1 <- ggplot(num_greedy_05_m30_pr2_11[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=30,pa=0.5,T,p1=0.03,p2=0.01,pr=2")+
 theme_bw()

p2 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_greedy_05_m30_pr2_11,size=0.1) +
  #ylim(0,4)+
  xlab("Time period")+
  labs(title = "Greedy - m=30,pa=0.5,T,p1=0.03,p2=0.01,pr=2") +
  theme_bw()


#-------------------------------------------------------------------------------
```

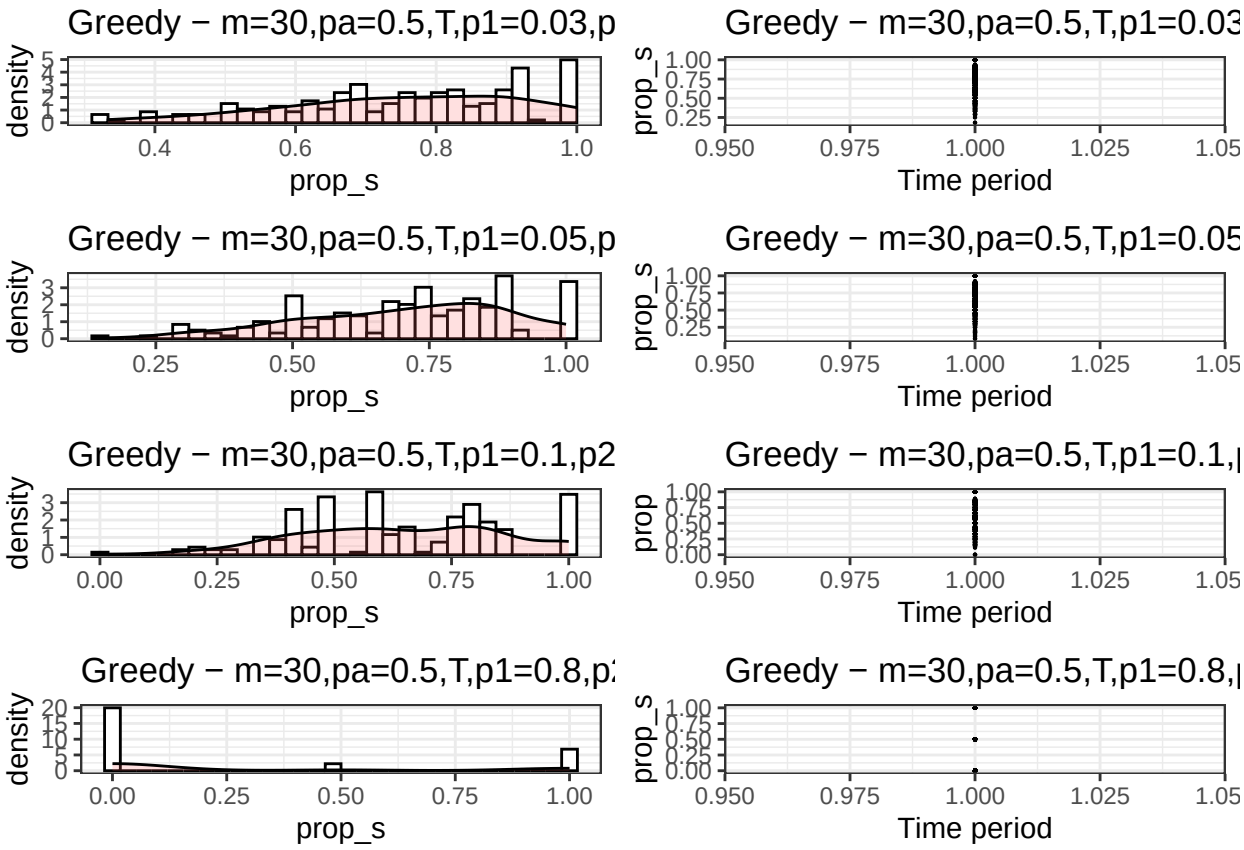
```
p3 <- ggplot(num_greedy_05_m30_pr2_12[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=30,pa=0.5,T,p1=0.05,p2=0.01,pr=2")+
 theme_bw()

p4 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_greedy_05_m30_pr2_12,size=0.1) +
  #ylim(0,5)+
  xlab("Time period")+
  labs(title = "Greedy - m=30,pa=0.5,T,p1=0.05,p2=0.01,pr=2") +
  theme_bw()


#------------------------------------------------------------------------------

p5 <- ggplot(num_greedy_05_m30_pr2_13[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=30,pa=0.5,T,p1=0.1,p2=0.01,pr=2")+
 theme_bw()

p6 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_greedy_05_m30_pr2_13,size=0.1) +
  xlab("Time period")+
  ylab("prop")+
  #ylim(0,6)+
  labs(title = "Greedy - m=30,pa=0.5,T,p1=0.1,p2=0.01,pr=2") +
  theme_bw()


#------------------------------------------------------------------------------

p7 <- ggplot(num_greedy_05_m30_pr2_14[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=30,pa=0.5,T,p1=0.8,p2=0.01,pr=2")+
 theme_bw()


p8 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_greedy_05_m30_pr2_14,size=0.1) +
  xlab("Time period")+
  labs(title = "Greedy - m=30,pa=0.5,T,p1=0.8,p2=0.01,pr=2") +
  theme_bw()


grid.arrange(p1,p2,p3,p4,p5,p6,p7,p8,nrow=4,ncol=2 )
```

2.1.1 market size m = 30

```
## Warning: The dot-dot notation (`..density..`) was deprecated in ggplot2 3.4.0.
## i Please use `after_stat(density)` instead.
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing non-finite values (`stat_bin()`).
## Warning: Removed 1 rows containing non-finite values (`stat_density()`).
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 44 rows containing non-finite values (`stat_bin()`).
## Warning: Removed 44 rows containing non-finite values (`stat_density()`).
## Warning: Removed 225 rows containing missing values (`geom_point()`).
```



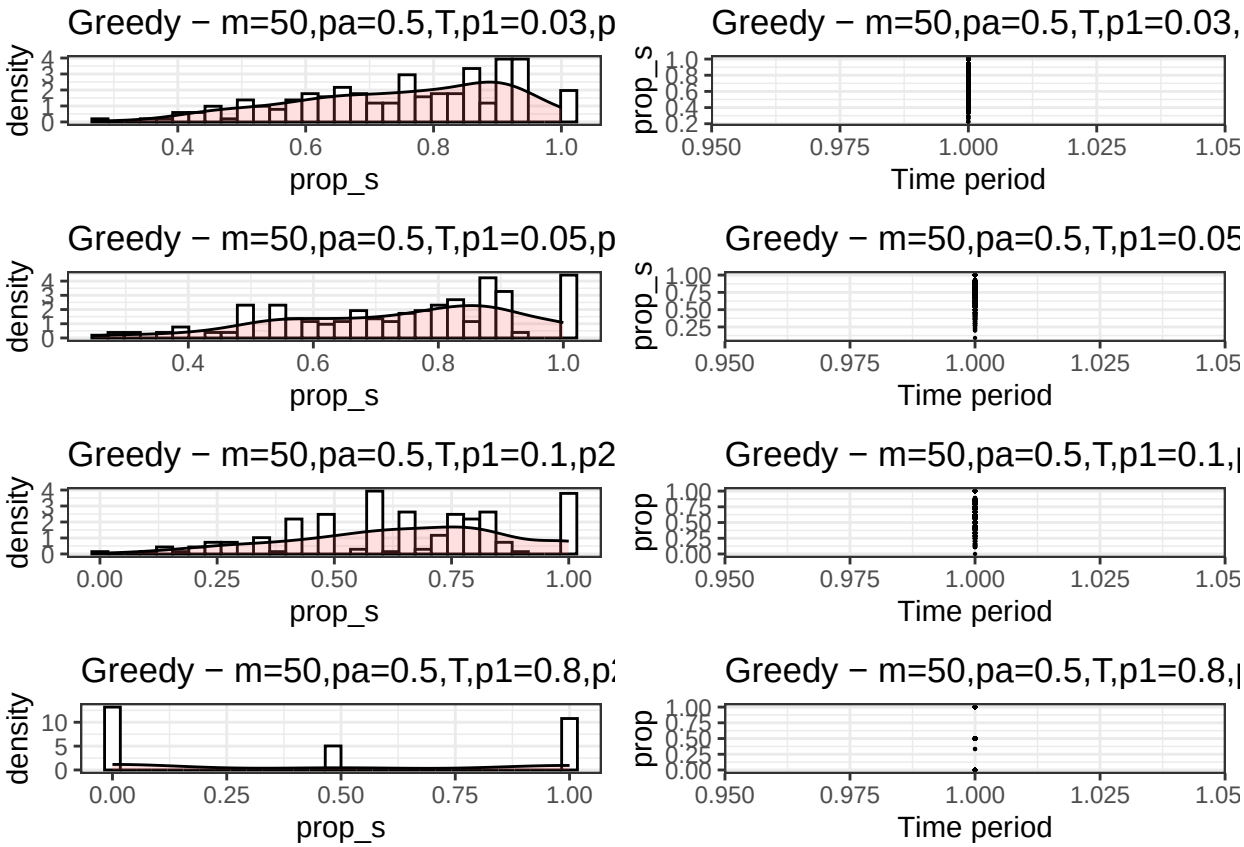

```
p1 <- ggplot(num_greedy_05_m50_pr2_11[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=50,pa=0.5,T,p1=0.03,p2=0.01,pr=2")+
 theme_bw()

p2 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_greedy_05_m50_pr2_11,size=0.1) +
```

```
  #ylim(0,4)+
  xlab("Time period")+
  labs(title = "Greedy - m=50,pa=0.5,T,p1=0.03,p2=0.01,pr=2") +
  theme_bw()


#-------------------------------------------------------------------------------

p3 <- ggplot(num_greedy_05_m50_pr2_12[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=50,pa=0.5,T,p1=0.05,p2=0.01,pr=2")+
 theme_bw()

p4 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_greedy_05_m50_pr2_12,size=0.1) +
  #ylim(0,5)+
  xlab("Time period")+
  labs(title = "Greedy - m=50,pa=0.5,T,p1=0.05,p2=0.01,pr=2") +
  theme_bw()


#-------------------------------------------------------------------------------

p5 <- ggplot(num_greedy_05_m50_pr2_13[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=50,pa=0.5,T,p1=0.1,p2=0.01,pr=2")+
 theme_bw()

p6 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_greedy_05_m50_pr2_13,size=0.1) +
  xlab("Time period")+
  ylab("prop")+
  labs(title = "Greedy - m=50,pa=0.5,T,p1=0.1,p2=0.01,pr=2") +
  theme_bw()


#-------------------------------------------------------------------------------

p7 <- ggplot(num_greedy_05_m50_pr2_14[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=50,pa=0.5,T,p1=0.8,p2=0.01,pr=2")+
 theme_bw()

p8 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_greedy_05_m50_pr2_14,size=0.1) +
  xlab("Time period")+
  ylab("prop")+
  labs(title = "Greedy - m=50,pa=0.5,T,p1=0.8,p2=0.01,pr=2") +
  theme_bw()
```

```
grid.arrange(p1,p2,p3,p4,p5,p6,p7,p8,nrow=4,ncol=2 )
```

2.1.2 market size m = 50

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 2 rows containing non-finite values (`stat_bin()`).
## Warning: Removed 2 rows containing non-finite values (`stat_density()`).
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 80 rows containing non-finite values (`stat_bin()`).
## Warning: Removed 80 rows containing non-finite values (`stat_density()`).
## Warning: Removed 213 rows containing missing values (`geom_point()`).
```



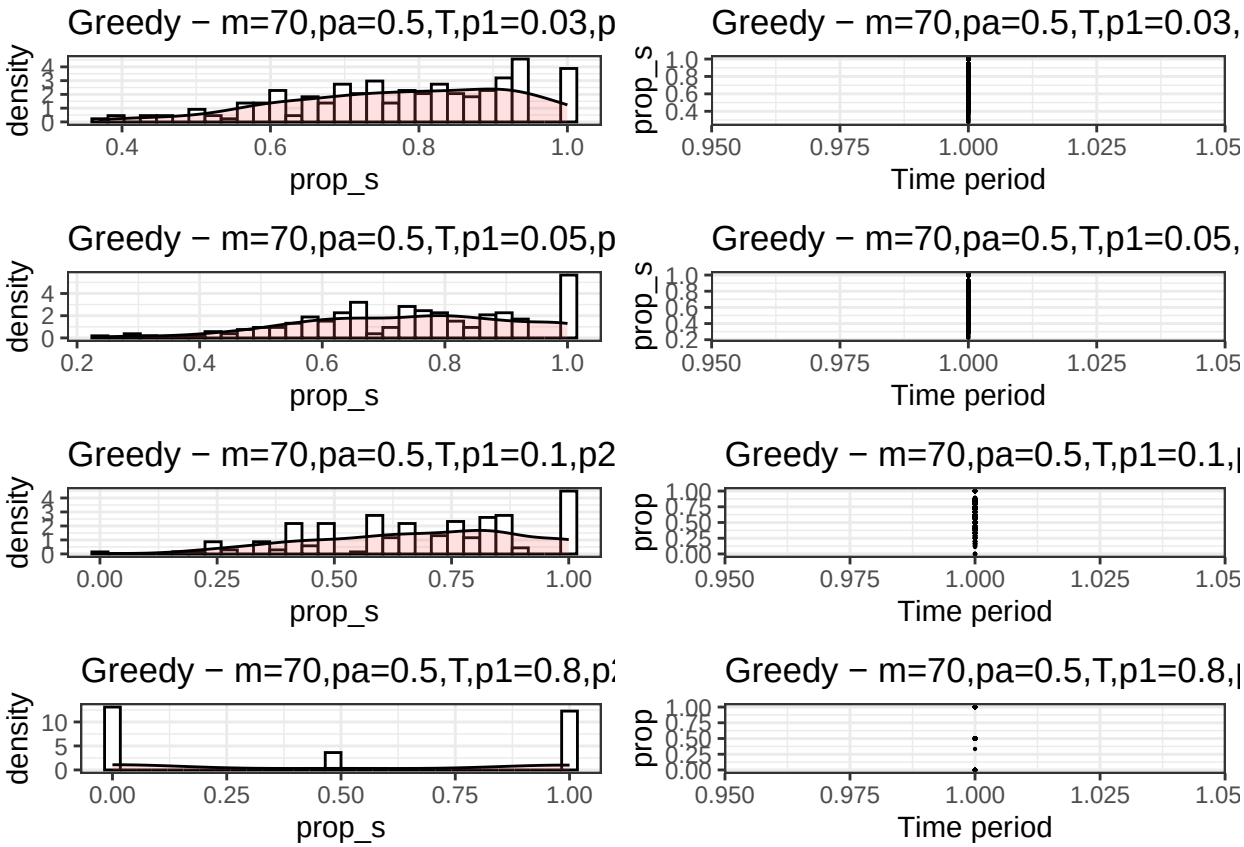

```
p1 <- ggplot(num_greedy_05_m70_pr2_11[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
```

```
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=70,pa=0.5,T,p1=0.03,p2=0.01,pr=2")+
 theme_bw()

p2 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_greedy_05_m70_pr2_11,size=0.1) +
  #ylim(0,4)+
  xlab("Time period")+
  labs(title = "Greedy - m=70,pa=0.5,T,p1=0.03,p2=0.01,pr=2") +
  theme_bw()


#-------------------------------------------------------------------------------

p3 <- ggplot(num_greedy_05_m70_pr2_12[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=70,pa=0.5,T,p1=0.05,p2=0.01,pr=2")+
 theme_bw()

p4 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_greedy_05_m70_pr2_12,size=0.1) +
  #ylim(0,5)+
  xlab("Time period")+
  labs(title = "Greedy - m=70,pa=0.5,T,p1=0.05,p2=0.01,pr=2") +
  theme_bw()

#-------------------------------------------------------------------------------

p5 <- ggplot(num_greedy_05_m70_pr2_13[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=70,pa=0.5,T,p1=0.1,p2=0.01,pr=2")+
 theme_bw()

p6 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_greedy_05_m70_pr2_13,size=0.1) +
  ylab("prop")+
  xlab("Time period")+
  labs(title = "Greedy - m=70,pa=0.5,T,p1=0.1,p2=0.01,pr=2") +
  theme_bw()

#-------------------------------------------------------------------------------

p7 <- ggplot(num_greedy_05_m70_pr2_14[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=70,pa=0.5,T,p1=0.8,p2=0.01,pr=2")+
 theme_bw()

p8 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_greedy_05_m70_pr2_14,size=0.1) +
  xlab("Time period")+
```

```
  ylab("prop")+
  labs(title = "Greedy - m=70,pa=0.5,T,p1=0.8,p2=0.01,pr=2") +
  theme_bw()


grid.arrange(p1,p2,p3,p4,p5,p6,p7,p8,nrow=4,ncol=2 )
```

2.1.3 market size m = 70

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing non-finite values (`stat_bin()`).
## Warning: Removed 1 rows containing non-finite values (`stat_density()`).
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 97 rows containing non-finite values (`stat_bin()`).
## Warning: Removed 97 rows containing non-finite values (`stat_density()`).
## Warning: Removed 221 rows containing missing values (`geom_point()`).
```

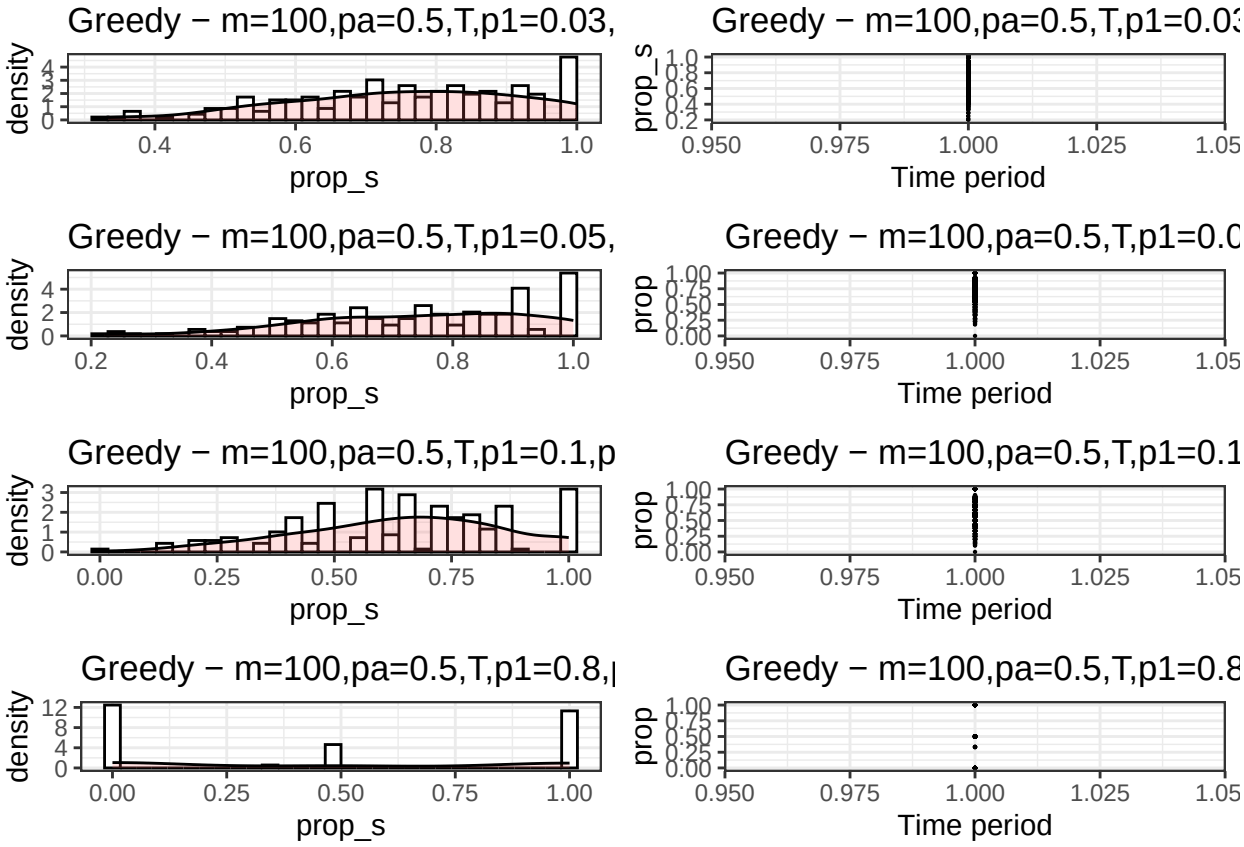


```
p1 <- ggplot(num_greedy_05_m100_pr2_11[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=100,pa=0.5,T,p1=0.03,p2=0.01,pr=2")+
 theme_bw()

p2 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_greedy_05_m100_pr2_11,size=0.1) +
  #ylim(0,4)+
  xlab("Time period")+
  labs(title = "Greedy - m=100,pa=0.5,T,p1=0.03,p2=0.01,pr=2") +
  theme_bw()

#-------------------------------------------------------------------------------

p3 <- ggplot(num_greedy_05_m100_pr2_12[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=100,pa=0.5,T,p1=0.05,p2=0.01,pr=2")+
 theme_bw()

p4 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_greedy_05_m100_pr2_12,size=0.1) +
  #ylim(0,5)+
  ylab("prop")+
  xlab("Time period")+
  labs(title = "Greedy - m=100,pa=0.5,T,p1=0.05,p2=0.01,pr=2") +
  theme_bw()

#-------------------------------------------------------------------------------

p5 <- ggplot(num_greedy_05_m100_pr2_13[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=100,pa=0.5,T,p1=0.1,p2=0.01,pr=2")+
 theme_bw()

p6 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_greedy_05_m100_pr2_13,size=0.1) +
  xlab("Time period")+
  ylab("prop")+
  labs(title = "Greedy - m=100,pa=0.5,T,p1=0.1,p2=0.01,pr=2") +
  theme_bw()

#-------------------------------------------------------------------------------

p7 <- ggplot(num_greedy_05_m100_pr2_14[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Greedy - m=100,pa=0.5,T,p1=0.8,p2=0.01,pr=2")+
 theme_bw()
```

```
p8 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_greedy_05_m100_pr2_14,size=0.1) +
  xlab("Time period")+
  ylab("prop")+
  labs(title = "Greedy - m=100,pa=0.5,T,p1=0.8,p2=0.01,pr=2") +
  theme_bw()


grid.arrange(p1,p2,p3,p4,p5,p6,p7,p8,nrow=4,ncol=2 )
```

2.1.4 market size m = 100

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 101 rows containing non-finite values (`stat_bin()`).
## Warning: Removed 101 rows containing non-finite values (`stat_density()`).
## Warning: Removed 218 rows containing missing values (`geom_point()`).
```



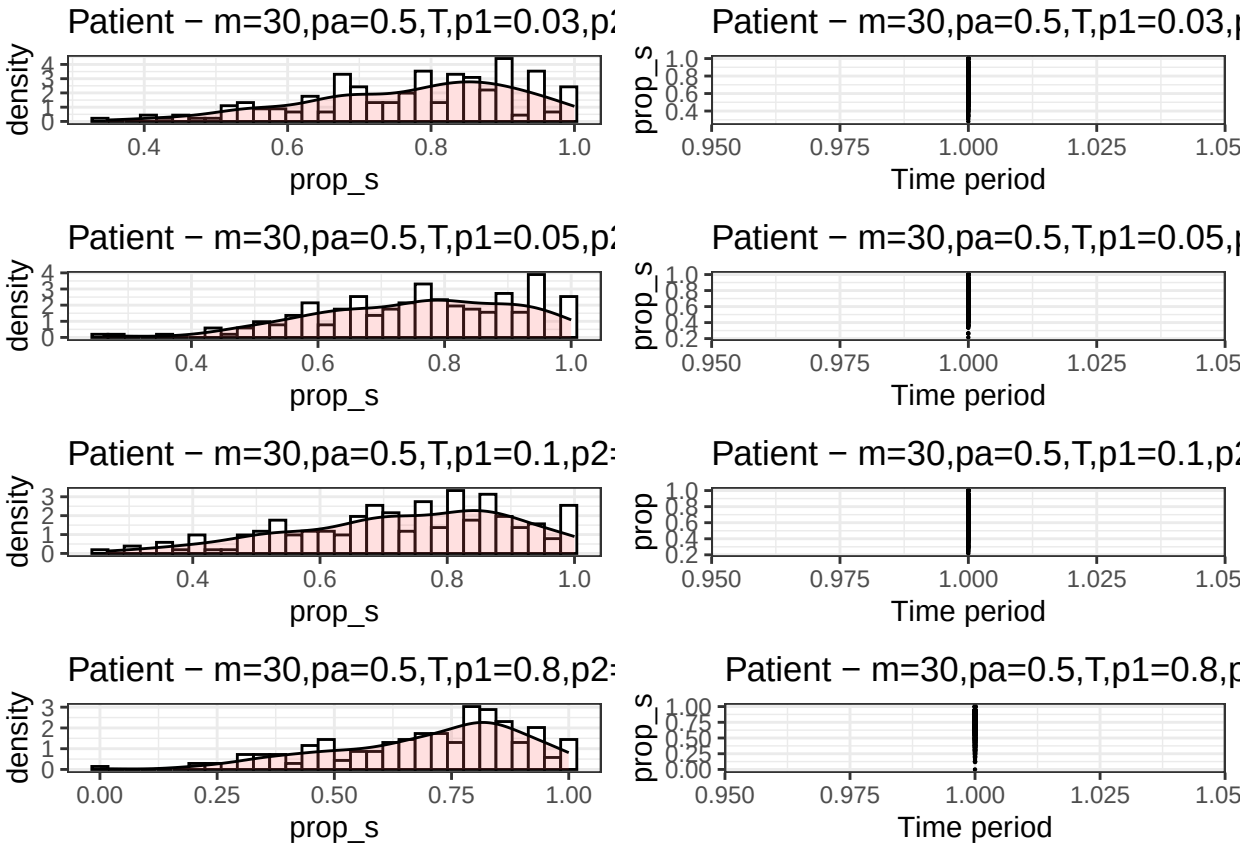
2.2 Patient algorithm plot

```
p1 <- ggplot(num_patient_05_m30_pr2_11[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=30,pa=0.5,T,p1=0.03,p2=0.01,pr=2")+
 theme_bw()

p2 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_patient_05_m30_pr2_11,size=0.1) +
  #ylim(0,4)+
  xlab("Time period")+
  labs(title = "Patient - m=30,pa=0.5,T,p1=0.03,p2=0.01,pr=2") +
  theme_bw()


#-------------------------------------------------------------------------------

p3 <- ggplot(num_patient_05_m30_pr2_12[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=30,pa=0.5,T,p1=0.05,p2=0.01,pr=2")+
 theme_bw()

p4 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_patient_05_m30_pr2_12,size=0.1) +
  #ylim(0,5)+
  xlab("Time period")+
  labs(title = "Patient - m=30,pa=0.5,T,p1=0.05,p2=0.01,pr=2") +
  theme_bw()


#-------------------------------------------------------------------------------

p5 <- ggplot(num_patient_05_m30_pr2_13[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=30,pa=0.5,T,p1=0.1,p2=0.01,pr=2")+
 theme_bw()

p6 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_patient_05_m30_pr2_13,size=0.1) +
  xlab("Time period")+
  ylab("prop")+
  #ylim(0,6)+
  labs(title = "Patient - m=30,pa=0.5,T,p1=0.1,p2=0.01,pr=2") +
  theme_bw()


#-------------------------------------------------------------------------------

p7 <- ggplot(num_patient_05_m30_pr2_14[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
```

```
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=30,pa=0.5,T,p1=0.8,p2=0.01,pr=2")+
 theme_bw()


p8 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_patient_05_m30_pr2_14,size=0.1) +
  xlab("Time period")+
  labs(title = "Patient - m=30,pa=0.5,T,p1=0.8,p2=0.01,pr=2") +
  theme_bw()


grid.arrange(p1,p2,p3,p4,p5,p6,p7,p8,nrow=4,ncol=2 )
```

2.2.1 market size m = 30

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
```

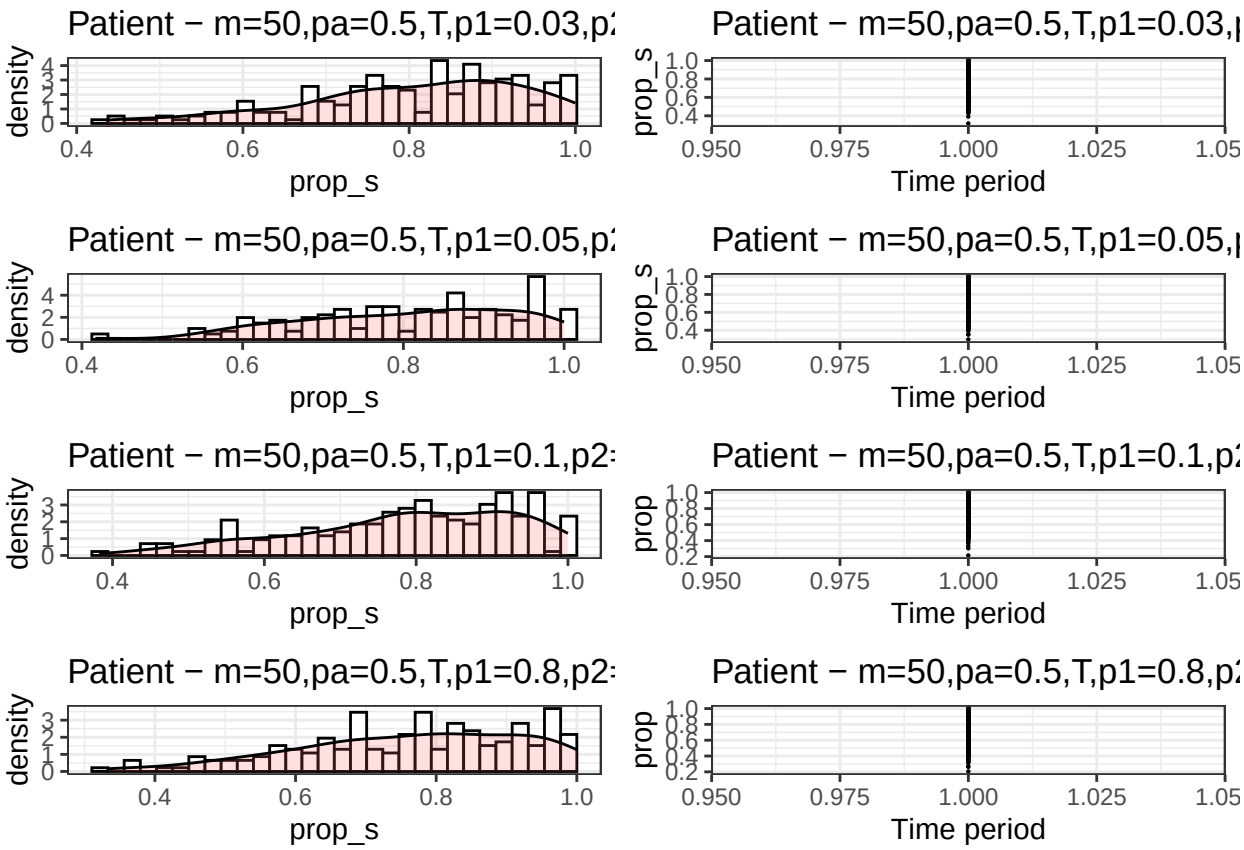



```
p1 <- ggplot(num_patient_05_m50_pr2_11[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=50,pa=0.5,T,p1=0.03,p2=0.01,pr=2")+
 theme_bw()

p2 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_patient_05_m50_pr2_11,size=0.1) +
  #ylim(0,4)+
  xlab("Time period")+
  labs(title = "Patient - m=50,pa=0.5,T,p1=0.03,p2=0.01,pr=2") +
  theme_bw()


#-------------------------------------------------------------------------------

p3 <- ggplot(num_patient_05_m50_pr2_12[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=50,pa=0.5,T,p1=0.05,p2=0.01,pr=2")+
 theme_bw()

p4 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_patient_05_m50_pr2_12,size=0.1) +
  #ylim(0,5)+
```

```
  xlab("Time period")+
  labs(title = "Patient - m=50,pa=0.5,T,p1=0.05,p2=0.01,pr=2") +
  theme_bw()


#------------------------------------------------------------------------------

p5 <- ggplot(num_patient_05_m50_pr2_13[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=50,pa=0.5,T,p1=0.1,p2=0.01,pr=2")+
 theme_bw()

p6 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_patient_05_m50_pr2_13,size=0.1) +
  xlab("Time period")+
  ylab("prop")+
  labs(title = "Patient - m=50,pa=0.5,T,p1=0.1,p2=0.01,pr=2") +
  theme_bw()


#------------------------------------------------------------------------------

p7 <- ggplot(num_patient_05_m50_pr2_14[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=50,pa=0.5,T,p1=0.8,p2=0.01,pr=2")+
 theme_bw()

p8 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_patient_05_m50_pr2_14,size=0.1) +
  xlab("Time period")+
  ylab("prop")+
  labs(title = "Patient - m=50,pa=0.5,T,p1=0.8,p2=0.01,pr=2") +
  theme_bw()


grid.arrange(p1,p2,p3,p4,p5,p6,p7,p8,nrow=4,ncol=2 )
```

2.2.2 market size m = 50

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
```

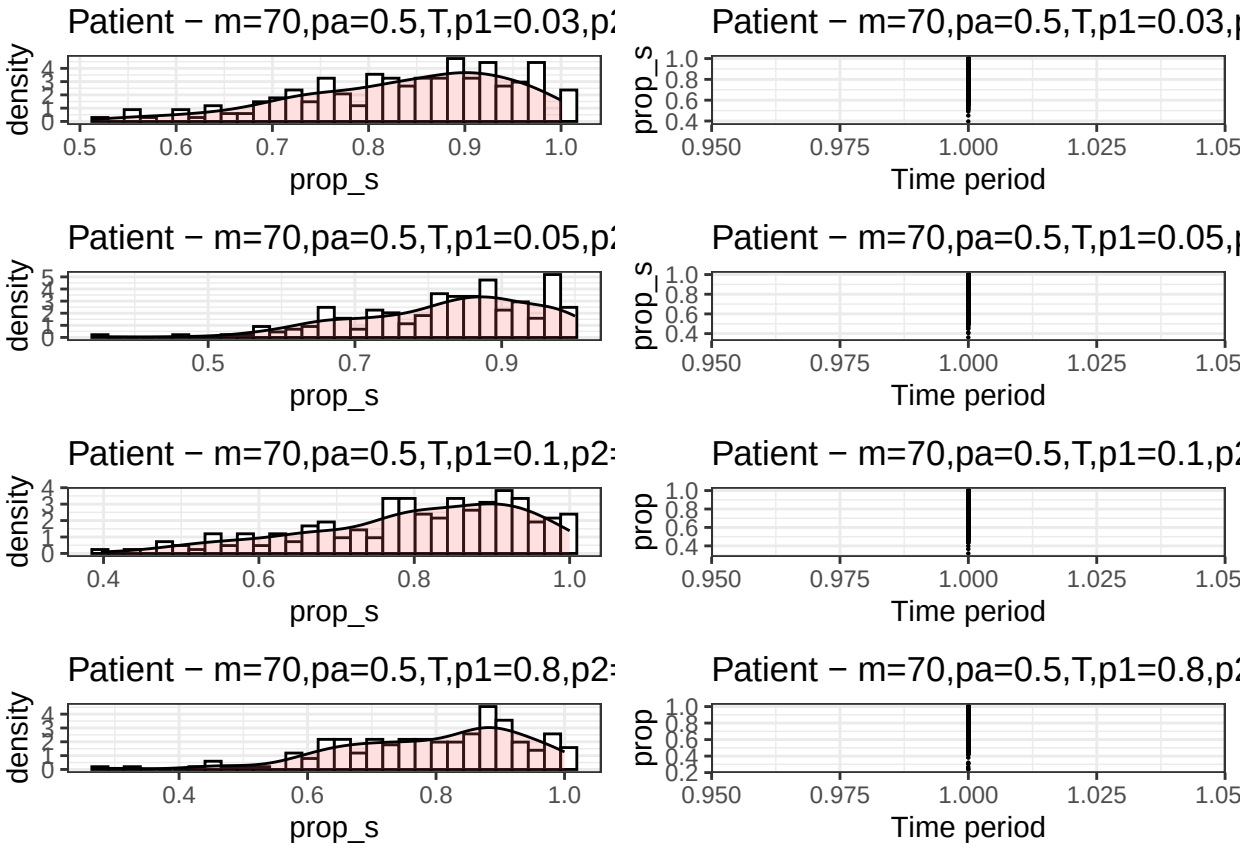



```
p1 <- ggplot(num_patient_05_m70_pr2_11[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=70,pa=0.5,T,p1=0.03,p2=0.01,pr=2")+
 theme_bw()

p2 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_patient_05_m70_pr2_11,size=0.1) +
  #ylim(0,4)+
  xlab("Time period")+
  labs(title = "Patient - m=70,pa=0.5,T,p1=0.03,p2=0.01,pr=2") +
  theme_bw()


#-------------------------------------------------------------------------------

p3 <- ggplot(num_patient_05_m70_pr2_12[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=70,pa=0.5,T,p1=0.05,p2=0.01,pr=2")+
 theme_bw()

p4 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_patient_05_m70_pr2_12,size=0.1) +
  #ylim(0,5)+
```

```
  xlab("Time period")+
  labs(title = "Patient - m=70,pa=0.5,T,p1=0.05,p2=0.01,pr=2") +
  theme_bw()

#-------------------------------------------------------------------------------

p5 <- ggplot(num_patient_05_m70_pr2_13[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=70,pa=0.5,T,p1=0.1,p2=0.01,pr=2")+
 theme_bw()

p6 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_patient_05_m70_pr2_13,size=0.1) +
  ylab("prop")+
  xlab("Time period")+
  labs(title = "Patient - m=70,pa=0.5,T,p1=0.1,p2=0.01,pr=2") +
  theme_bw()

#-------------------------------------------------------------------------------

p7 <- ggplot(num_patient_05_m70_pr2_14[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=70,pa=0.5,T,p1=0.8,p2=0.01,pr=2")+
 theme_bw()

p8 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_patient_05_m70_pr2_14,size=0.1) +
  xlab("Time period")+
  ylab("prop")+
  labs(title = "Patient - m=70,pa=0.5,T,p1=0.8,p2=0.01,pr=2") +
  theme_bw()


grid.arrange(p1,p2,p3,p4,p5,p6,p7,p8,nrow=4,ncol=2 )
```

2.2.3 market size m = 70

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
```

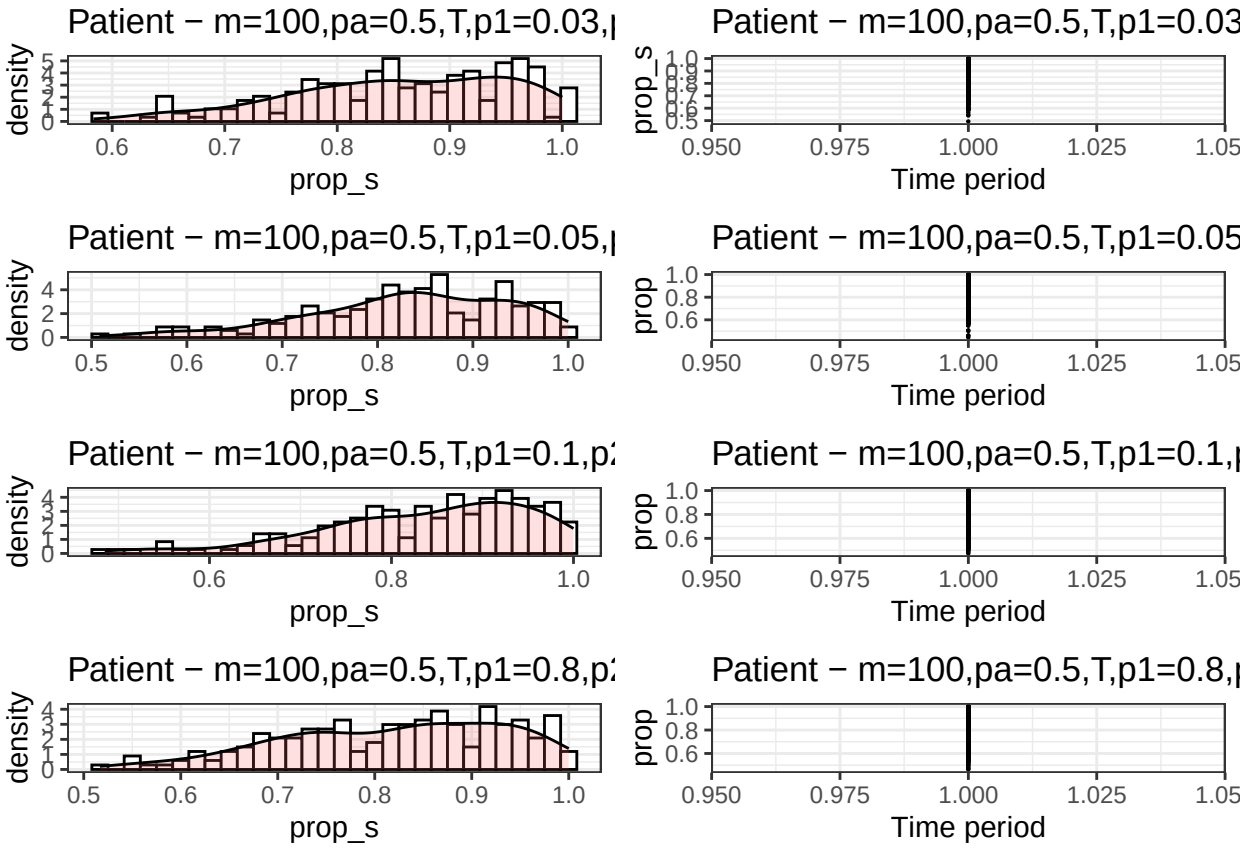



```
p1 <- ggplot(num_patient_05_m100_pr2_11[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=100,pa=0.5,T,p1=0.03,p2=0.01,pr=2")+
 theme_bw()

p2 <- ggplot(data = NULL, aes(x = c(1:T), y = prop_s)) +
  geom_point(data=num_patient_05_m100_pr2_11,size=0.1) +
  #ylim(0,4)+
  xlab("Time period")+
  labs(title = "Patient - m=100,pa=0.5,T,p1=0.03,p2=0.01,pr=2") +
  theme_bw()

#-------------------------------------------------------------------------------

p3 <- ggplot(num_patient_05_m100_pr2_12[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=100,pa=0.5,T,p1=0.05,p2=0.01,pr=2")+
 theme_bw()

p4 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_patient_05_m100_pr2_12,size=0.1) +
  #ylim(0,5)+
  ylab("prop")+
```

```
  xlab("Time period")+
  labs(title = "Patient - m=100,pa=0.5,T,p1=0.05,p2=0.01,pr=2") +
  theme_bw()

#-------------------------------------------------------------------------------

p5 <- ggplot(num_patient_05_m100_pr2_13[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=100,pa=0.5,T,p1=0.1,p2=0.01,pr=2")+
 theme_bw()

p6 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_patient_05_m100_pr2_13,size=0.1) +
  xlab("Time period")+
  ylab("prop")+
  labs(title = "Patient - m=100,pa=0.5,T,p1=0.1,p2=0.01,pr=2") +
  theme_bw()

#-------------------------------------------------------------------------------

p7 <- ggplot(num_patient_05_m100_pr2_14[800:1000,], aes(x=prop_s)) +
 geom_histogram(aes(y=..density..), colour="black", fill="white")+
 geom_density(alpha=.2, fill="#FF6666")+
 labs(title = "Patient - m=100,pa=0.5,T,p1=0.8,p2=0.01,pr=2")+
 theme_bw()

p8 <- ggplot(data = NULL, aes(x = c(1:T), y = as.double(prop_s))) +
  geom_point(data=num_patient_05_m100_pr2_14,size=0.1) +
  xlab("Time period")+
  ylab("prop")+
  labs(title = "Patient - m=100,pa=0.5,T,p1=0.8,p2=0.01,pr=2") +
  theme_bw()


grid.arrange(p1,p2,p3,p4,p5,p6,p7,p8,nrow=4,ncol=2)
```

2.2.4 market size m = 100

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 1 rows containing missing values (`geom_point()`).
```



Findings

- According to the plot of proportion of s, most of the distribution of proportion are all left-skewed distribution. In Patient algorithm, the density plots of proportion are all left-skewed distribution; while in Greedy algorithm, some of the density plot of the proportion could not be plotted due to many inf and 0 values, which indicates that one type agents would often vanish (become zero) under steady state. At this time, the dotplot of proportion would be shown as several horizontal lines, which means the value of proportion is among several fixed values. This case happens only when either compatible probability is rather large or newly coming agents arrive in the market in a higher rate.
- Based on these observations, we assume that the distribution of proportion is gamma distribution $\sim Beta(\alpha, \beta)$.

Fitted Beta distribution to s

Fit $Beta(\alpha, \beta)$ on the data.

```
# Greedy m=70, pa=0.5, p1=0.05, p2=0.01, pr=2
data <-num_greedy_05_m70_pr2_12$prop_s

exp_s<-mean(data[500:1000], na.rm = TRUE)
var_s<-var(data[500:1000],na.rm = TRUE)

alpha = exp_s*(exp_s*(1-exp_s)/var_s-1)
beta  = (1-exp_s)*(exp_s*(1-exp_s)/var_s-1)
```

```
alpha
```

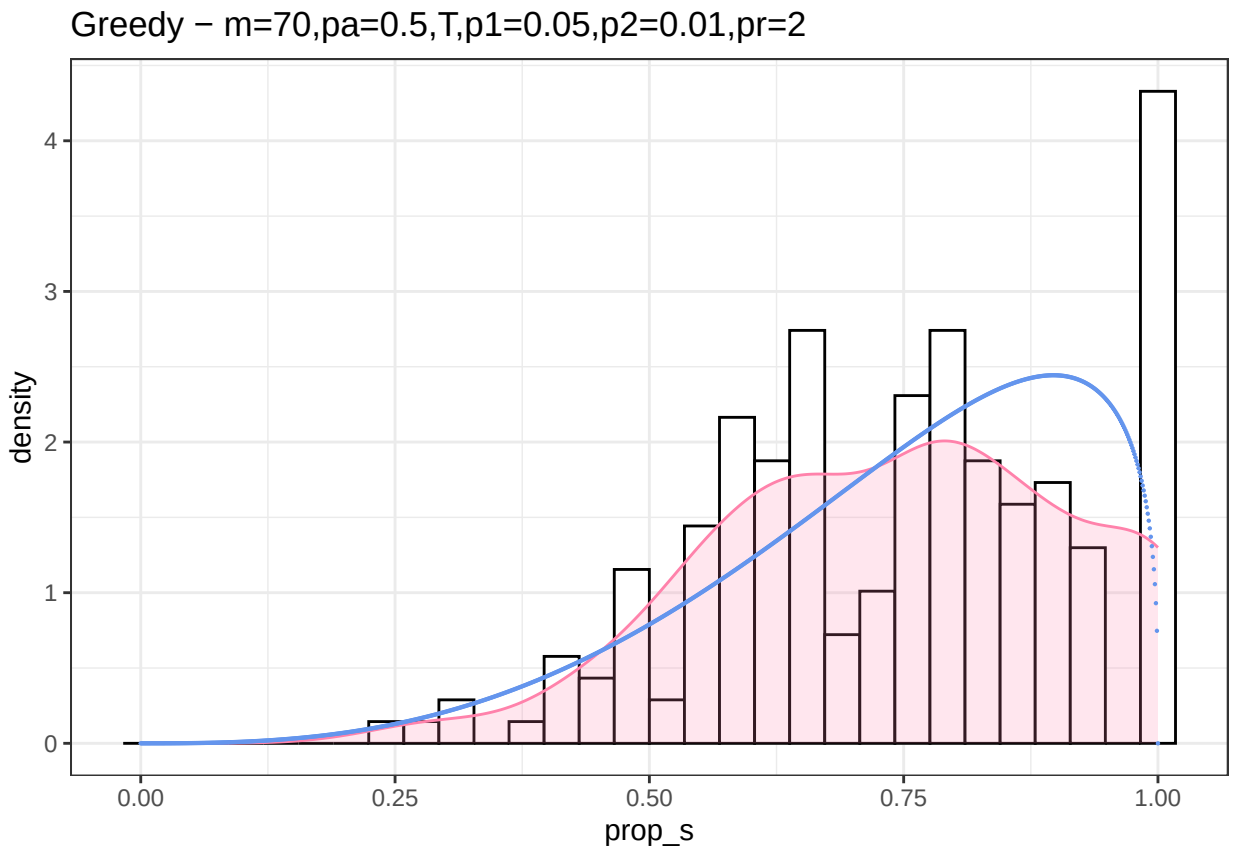
```
## [1] 3.801253
```

```
beta
```

```
## [1] 1.320809
```

```
# plot the beta distribution gamma(alpha, beta)

x <- seq(0,1,by=0.001)
fitted_val <- dbeta(x,shape1=alpha,shape2=beta)

fitted_data<-data.frame(x,fitted_val)


ggplot(num_greedy_05_m70_pr2_12[800:1000,], aes(x=prop_s)) +
  geom_histogram(aes(y=..density..), colour="black", fill="white")+
  geom_density(alpha=.2, colour="palevioletred1",fill="palevioletred1")+
  geom_point(data = fitted_data, aes(x = x, y = fitted_val), colour = "cornflowerblue",size=0.1,show.leg
  labs(title = "Greedy - m=70,pa=0.5,T,p1=0.05,p2=0.01,pr=2")+
  theme_bw()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
# Greedy m=100, pa=0.5, p1=0.1, p2=0.01, pr=2
data <-num_greedy_05_m100_pr2_13$prop_s
```

```r
exp_s<-mean(data[500:1000], na.rm = TRUE)
var_s<-var(data[500:1000],na.rm = TRUE)

alpha = exp_s*(exp_s*(1-exp_s)/var_s-1)
beta  = (1-exp_s)*(exp_s*(1-exp_s)/var_s-1)

alpha
```

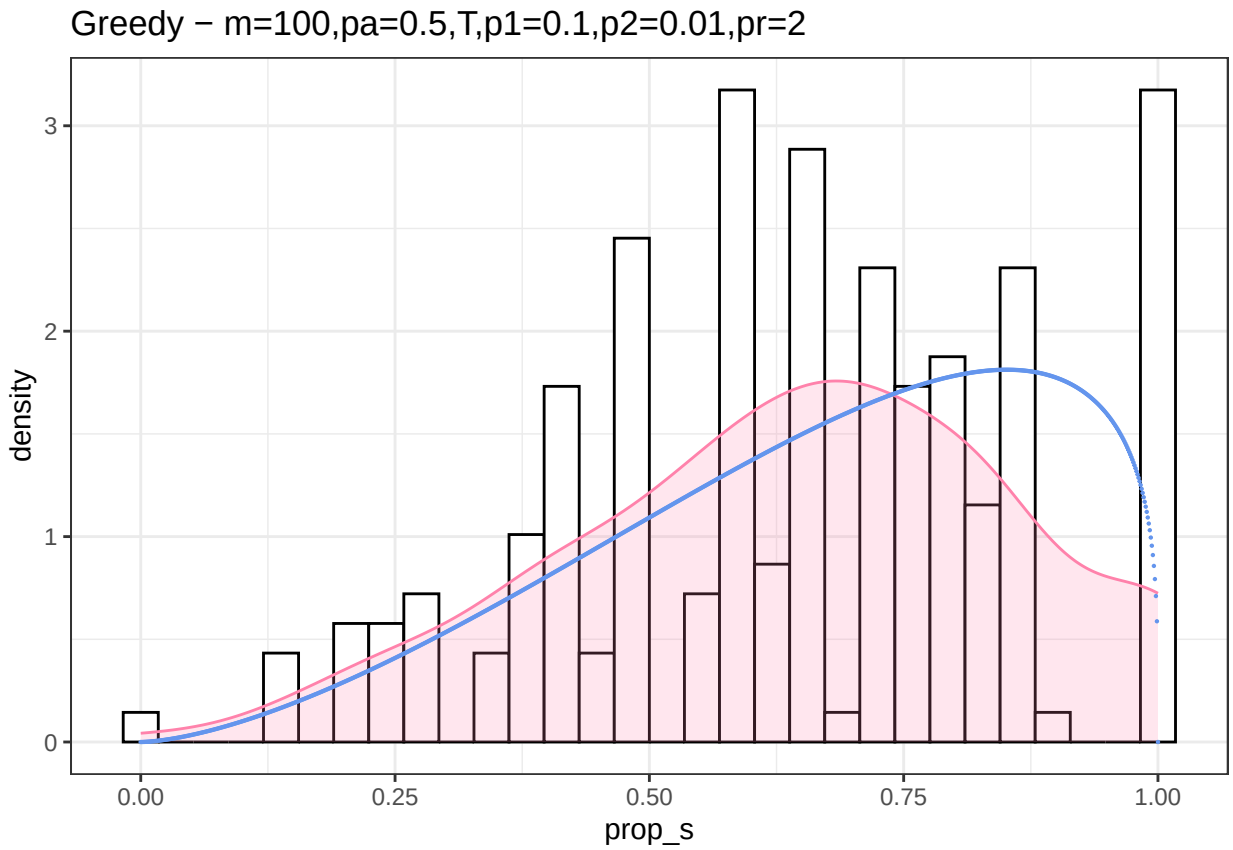
```
## [1] 2.57882
```

```r
beta
```

```
## [1] 1.275602
```

```r
# plot the beta distribution gamma(alpha, beta)

x <- seq(0,1,by=0.001)
fitted_val <- dbeta(x,shape1=alpha,shape2=beta)

fitted_data<-data.frame(x,fitted_val)


ggplot(num_greedy_05_m100_pr2_13[800:1000,], aes(x=prop_s)) +
  geom_histogram(aes(y=..density..), colour="black", fill="white")+
  geom_density(alpha=.2, colour="palevioletred1",fill="palevioletred1")+
  geom_point(data = fitted_data, aes(x = x, y = fitted_val), colour = "cornflowerblue",size=0.1,show.leg
  labs(title = "Greedy - m=100,pa=0.5,T,p1=0.1,p2=0.01,pr=2")+
  theme_bw()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
# Patient m=70, pa=0.5, p1=0.05, p2=0.01, pr=2
data <-num_patient_05_m70_pr2_12$prop_s

exp_s<-mean(data[500:1000], na.rm = TRUE)
var_s<-var(data[500:1000],na.rm = TRUE)

alpha = exp_s*(exp_s*(1-exp_s)/var_s-1)
beta  = (1-exp_s)*(exp_s*(1-exp_s)/var_s-1)

alpha
```

```
## [1] 7.101472
```

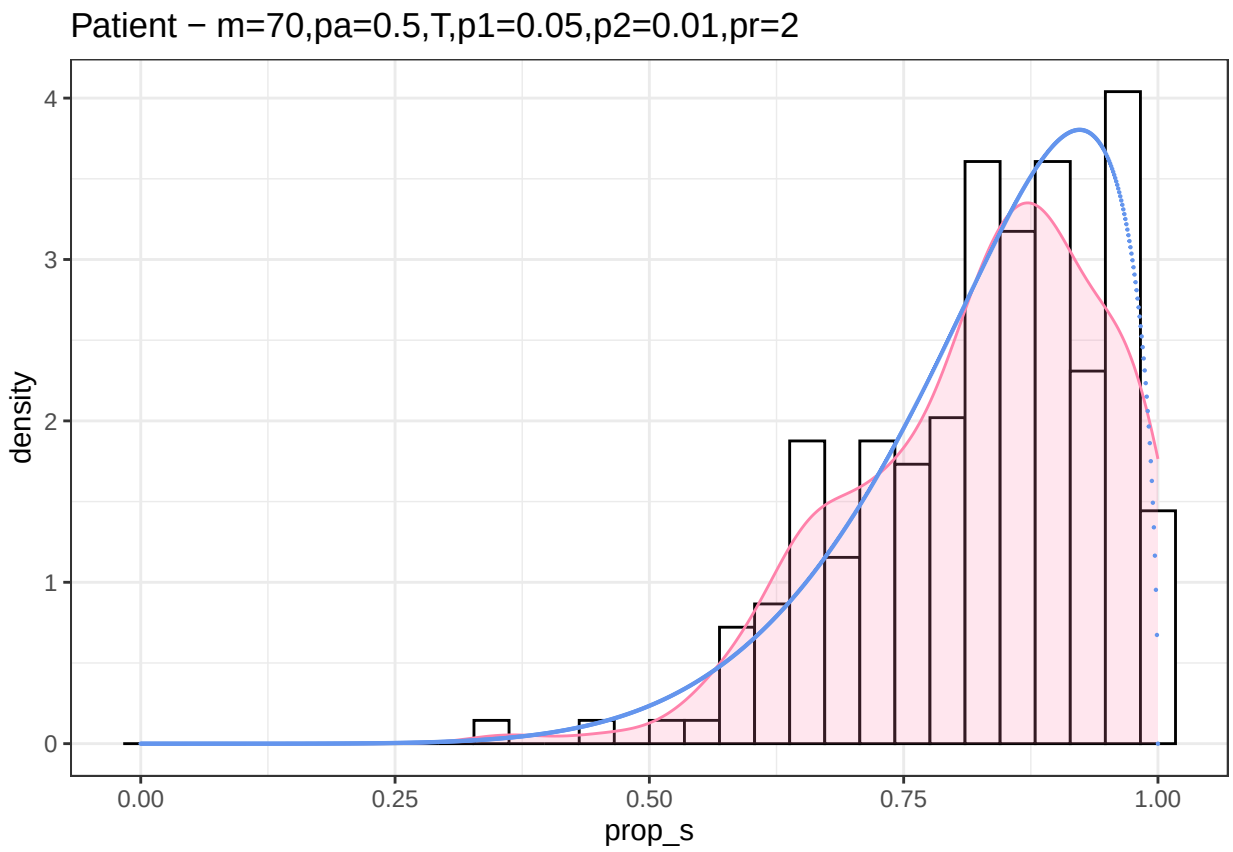
```
beta
```

```
## [1] 1.509657
```

```
# plot the beta distribution gamma(alpha, beta)

x <- seq(0,1,by=0.001)
fitted_val <- dbeta(x,shape1=alpha,shape2=beta)

fitted_data<-data.frame(x,fitted_val)


ggplot(num_patient_05_m70_pr2_12[800:1000,], aes(x=prop_s)) +
  geom_histogram(aes(y=..density..), colour="black", fill="white")+
  geom_density(alpha=.2, colour="palevioletred1",fill="palevioletred1")+
  geom_point(data = fitted_data, aes(x = x, y = fitted_val), colour = "cornflowerblue",size=0.1,show.leg
```

```
  labs(title = "Patient - m=70,pa=0.5,T,p1=0.05,p2=0.01,pr=2")+
  theme_bw()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
# Patient m=100, pa=0.5, p1=0.1, p2=0.01, pr=2
data <-num_patient_05_m100_pr2_13$prop_s

exp_s<-mean(data[500:1000], na.rm = TRUE)
var_s<-var(data[500:1000],na.rm = TRUE)

alpha = exp_s*(exp_s*(1-exp_s)/var_s-1)
beta  = (1-exp_s)*(exp_s*(1-exp_s)/var_s-1)

alpha
```

```
## [1] 7.665775
```

```
beta
```

```
## [1] 1.497327
```

```
# plot the beta distribution gamma(alpha, beta)

x <- seq(0,1,by=0.001)
fitted_val <- dbeta(x,shape1=alpha,shape2=beta)

fitted_data<-data.frame(x,fitted_val)
```

```
ggplot(num_patient_05_m100_pr2_13[800:1000,], aes(x=prop_s)) +
  geom_histogram(aes(y=..density..), colour="black", fill="white")+
  geom_density(alpha=.2, colour="palevioletred1",fill="palevioletred1")+
  geom_point(data = fitted_data, aes(x = x, y = fitted_val), colour = "cornflowerblue",size=0.1,show.leg
  labs(title = "Patient - m=100,pa=0.5,T,p1=0.1,p2=0.01,pr=2")+
  theme_bw()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```