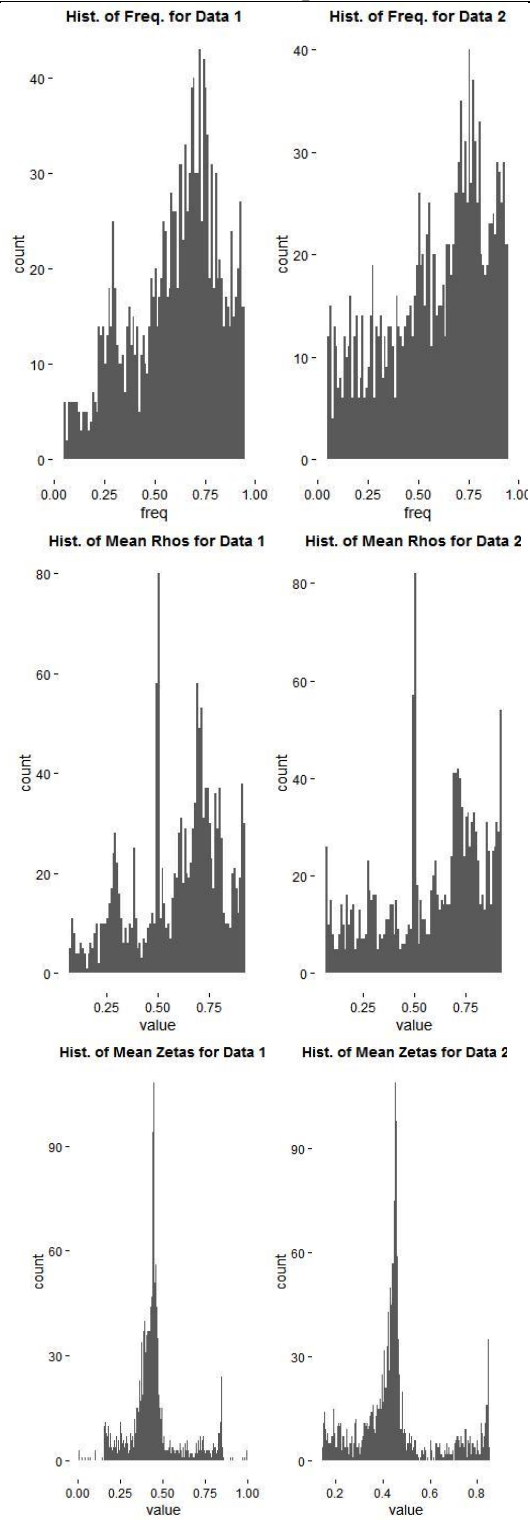


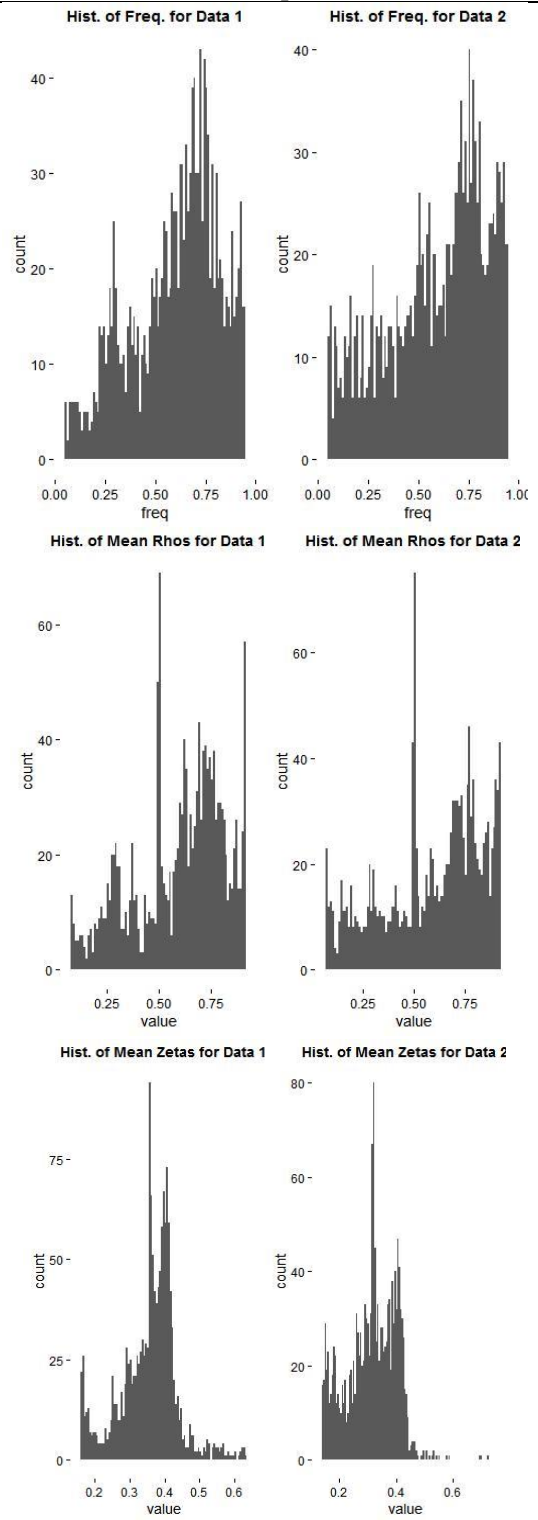
Comparing Real Data with the MH Split and without

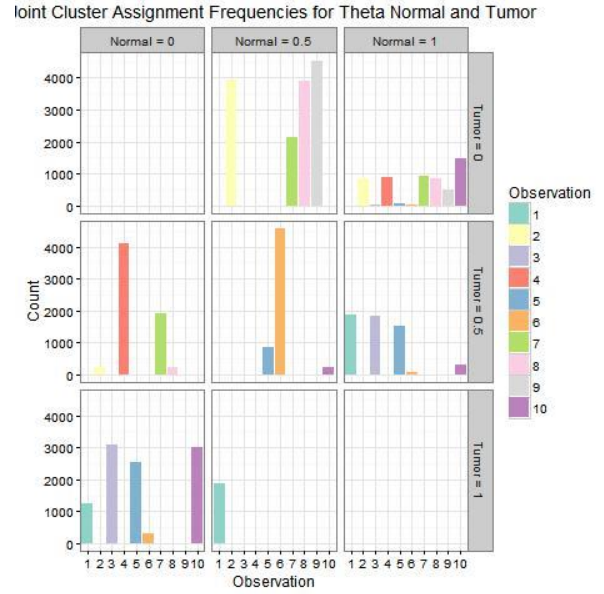
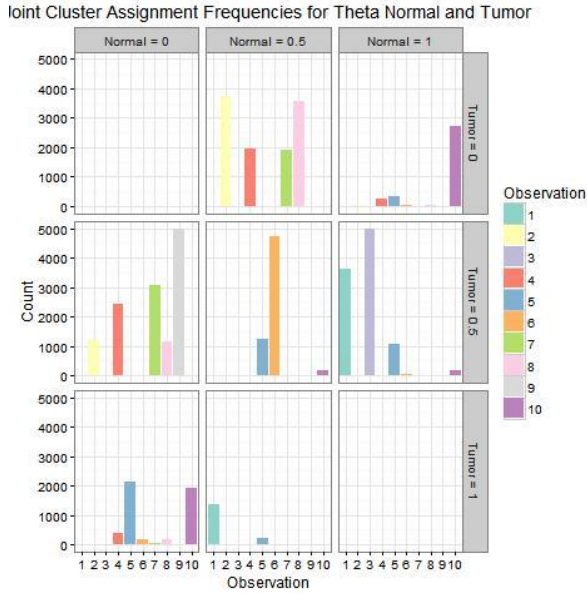
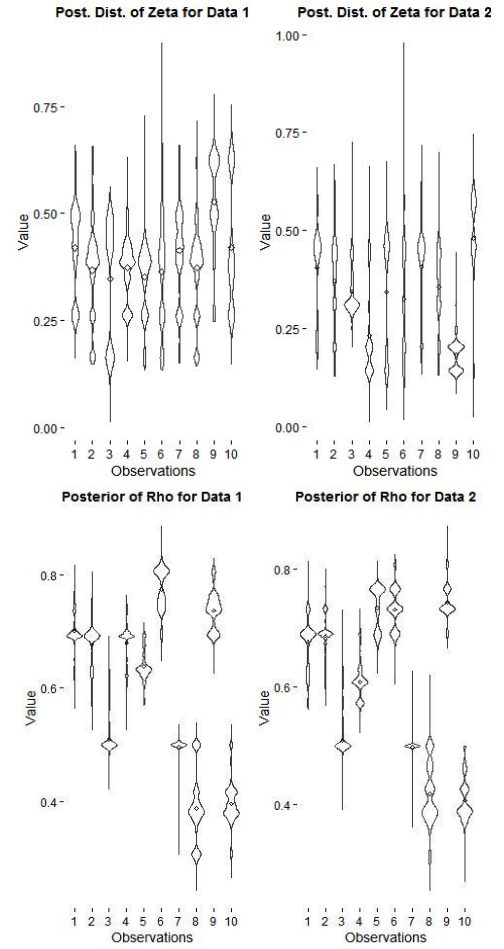
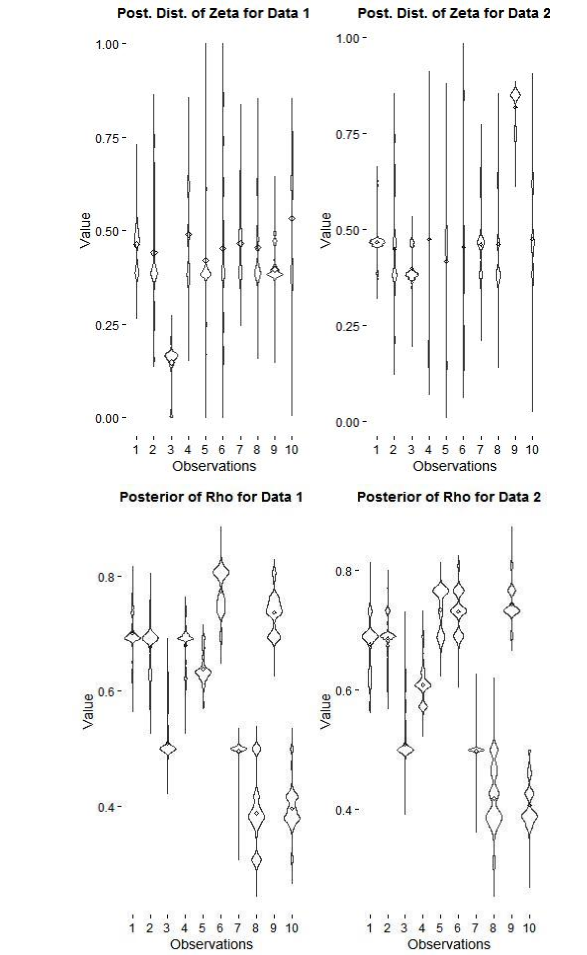
Here are several results of the two algorithms

No MH Split



MH Split





If we look at observation number 2, we can see that the algorithm explored a third mode. If we look at observation 3 from the above, before split, the algorithm only explored one mode $\theta = (\theta_N, \theta_T) = (1, 0.5)$. On the other hand, after the split, the algorithm explores a second mode which corresponds to $\theta = (0, 1)$. Here is the estimated value for data 1 and 2 respectively.

With split:

	theta_n_hat	theta_t_hat	mean_zeta	mean_rho	real_rho
3	0.0	1.0	0.3464	0.5172	0.54
3	0.0	1.0	0.3442	0.5168	0.69

$$\rho_1 = 0.35 \times 0 + 0.65 \times 1 = 0.65$$

$$\rho_2 = 0.34 \times 0 + 0.66 \times 1 = 0.66$$

These ρ could also be a potential result.

With no split:

	theta_n_hat	theta_t_hat	mean_zeta	mean_rho	real_rho
3	1.0	0.5	0.1482	0.5097	0.54
3	1.0	0.5	0.3982	0.5097	0.69

$$\rho_1 = 0.15 \times 1 + 0.85 \times 0.5 = 0.58$$

$$\rho_2 = 0.4 \times 1 + 0.6 \times 0.5 = 0.7$$

So this is pretty close to the real rho.

Here is the real data:

	trials	mixDat	freq
947	344	186	0.5406977
947	317	220	0.69400631

The next observation, I wanted to look at is observation 9:

With split:

	theta_n_hat	theta_t_hat	mean_zeta	mean_rho	real_rho
9	0.5	0.0	0.5271	0.7462	0.31
9	0.5	0.0	0.1831	0.7567	0.09

$$\rho_1 = 0.53 \times 0.5 + 0.47 \times 0 = 0.27$$

$$\rho_2 = 0.18 \times 0.5 + 0.82 \times 0 = 0.09$$

These ρ could also be a potential result but I am not sure why the mean_rho is significantly different to the real rho.

With no split:

	theta_n_hat	theta_t_hat	mean_zeta	mean_rho	real_rho
9	0.0	0.5	0.3975	0.7368	0.31
9	0.0	0.5	0.8179	0.7442	0.09

$$\rho_1 = 0.40 \times 0 + 0.60 \times 0.5 = 0.30$$

$$\rho_2 = 0.82 \times 0 + 0.18 \times 0.5 = 0.09$$

So this is pretty close to the real rho and it definitely show the exploration of more modes.

Here is the real data:

	trials	mixDat	freq
1032	121	37	0.3057851
1032	144	13	0.09027778

The last observation I looked at was observation 1:

With split:

	theta_n_hat	theta_t_hat	mean_zeta	mean_rho	real_rho
1	1.0	0.5	0.4192	0.6967	0.73
1	1.0	0.5	0.4055	0.6650	0.76

$$\rho_1 = 0.42 \times 1 + 0.58 \times 0.5 = 0.71$$

$$\rho_2 = 0.41 \times 1 + 0.59 \times 0.5 = 0.71$$

With no split:

	theta_n_hat	theta_t_hat	mean_zeta	mean_rho	real_rho
1	1.0	0.5	0.4621	0.7008	0.73
1	1.0	0.5	0.4669	0.6766	0.76

$$\rho_1 = 0.46 \times 1 + 0.66 \times 0.5 = 0.79$$

$$\rho_2 = 0.47 \times 1 + 0.53 \times 0.5 = 0.73$$

But the alternative here is $\theta = (0.5, 1)$ with I think equal number of visit to this mode. If this is the case, then the $\zeta_1 = 0.54$ and $\zeta_2 = 0.48$. This would be an alternative solution to the equation.

Here is the real data:

	trials	mixDat	freq
177	304	223	0.7335526
177	261	198	0.75862069

Here are the real data for all 10:

```
> data1[test_obs1,]
      trials mixDat      freq cluster ucl_n ucl_t
177      304    223 0.7335526     839     3     3
968      120     23 0.1916667     858     2     1
947      344    186 0.5406977    1196     1     2
969      148     45 0.3040541      32     3     1
1337     123     78 0.6341463     701     2     2
994      116     58 0.5000000     980     1     3
15       110     26 0.2363636    1466     2     2
361      122     24 0.1967213     618     1     2
1032     121     37 0.3057851     598     2     2
796      103     70 0.6796117      86     2     2
```

```
> data2[test_obs1,]
      trials mixDat      freq cluster ucl_n ucl_t
177      261    198 0.75862069      674     3     3
968      142     28 0.19718310     1059     2     1
947      317    220 0.69400631     1506     1     2
969      146     62 0.42465753      602     3     1
1337     157     80 0.50955414     1556     2     2
994      135     75 0.55555556      333     1     3
15       103     25 0.24271845      600     2     2
361      145     27 0.18620690     1261     1     2
1032     144     13 0.09027778      427     2     2
796      123     53 0.43089431      609     2     2
```

Here are the estimates with Split:

```
> print(est_data1)
      theta_n_hat theta_t_hat mean_zeta_1 mean_rho_1 real_rho_1
1             1.0          0.5      0.4192      0.6967      0.73
2             0.5          0.0      0.3671      0.6619      0.19
3             0.0          1.0      0.3464      0.5172      0.54
4             0.0          0.5      0.3732      0.6718      0.30
5             1.0          0.5      0.3516      0.6372      0.63
6             0.5          0.5      0.3632      0.7789      0.50
7             0.5          0.0      0.4128      0.4938      0.24
8             0.5          0.0      0.3723      0.3784      0.20
9             0.5          0.0      0.5271      0.7462      0.31
10            0.0          1.0      0.4203      0.3924      0.68

> print(est_data2)
      theta_n_hat theta_t_hat mean_zeta_2 mean_rho_2 real_rho_2
1             1.0          0.5      0.4055      0.6650      0.76
2             0.5          0.0      0.3714      0.6780      0.20
3             0.0          1.0      0.3442      0.5168      0.69
4             0.0          0.5      0.2311      0.6069      0.42
5             1.0          0.5      0.3426      0.7160      0.51
6             0.5          0.5      0.3243      0.7424      0.56
7             0.5          0.0      0.4081      0.4964      0.24
8             0.5          0.0      0.3560      0.4118      0.19
9             0.5          0.0      0.1831      0.7567      0.09
10            0.0          1.0      0.4792      0.4101      0.43
```

Here are the estimates with no Split:

```
> print(est_data1)
      theta_n_hat theta_t_hat mean_zeta_1 mean_rho_1 real_rho_1
1             1.0          0.5      0.4621      0.7008      0.73
2             0.5          0.0      0.4408      0.6770      0.19
3             1.0          0.5      0.1482      0.5097      0.54
4             0.5          0.0      0.4900      0.6796      0.30
5             1.0          0.5      0.4198      0.6385      0.63
6             0.5          0.5      0.4525      0.7744      0.50
```

```

7      0.0      0.0      0.4661      0.4960      0.24
8      0.5      0.0      0.4539      0.3880      0.20
9      0.0      0.5      0.3975      0.7368      0.31
10     1.0      0.0      0.5332      0.3962      0.68
> print(est_data2)
      theta_n_hat theta_t_hat mean_zeta_2 mean_rho_2 real_rho_2
1      1.0      0.5      0.4669      0.6766      0.76
2      0.5      0.0      0.4490      0.6872      0.20
3      1.0      0.5      0.3982      0.5097      0.69
4      0.5      0.0      0.4749      0.6090      0.42
5      1.0      0.5      0.4169      0.7345      0.51
6      0.5      0.5      0.4541      0.7317      0.56
7      0.0      0.0      0.4602      0.4979      0.24
8      0.5      0.0      0.4621      0.4186      0.19
9      0.0      0.5      0.8179      0.7442      0.09
10     1.0      0.0      0.4749      0.4077      0.43

```