Sequential M-MED

Kristyn Pantoja

10/04/2019

M-MED with Data

An Example

Bn

MSE(Bn)

Posterior Probabilities of Hypotheses

M-MED with Data

How to Make M-MED Sequential, using data:

- To generate M-MED, call it D, given two hypothesized models H_0 and H_1, we try to choose design points that minimize the total potential energy, where the charge function of point $x \in D$ is given by $f_D = W(\phi_{0,x}, \phi_{1,x})$.
- Without data, $\phi_{i,x}$, $i \in \{0, 1\}$ is the density of the marginal distribution of y, given hypothesis i's model
  - i.e. $\phi_{i,x}$ is the pdf of $N(\mathbf{x}^T \mu_i, \sigma_\varepsilon^2 + \mathbf{x}^T V_0 \mathbf{x})$
  - $\mathbf{x}$ is the row of the design matrix $\mathbf{X}$ corresponding to point x.
- To incorporate data in choosing the design, $\phi_{i,x}$ is the posterior predictive distribution of $\tilde{y}$ (unknown) at untested point $\tilde{x}$, given previous data (X, y).
  - We choose the design points for X by using the M-MED method without data.

Posterior Predictive Distribution of y

$[\tilde{y}|\tilde{x}, X, y, \sigma_\varepsilon^2, H_i, V_i]$ for brevity, call it $\tilde{y}|y$

$$f(\tilde{y}|y) = \int f(\tilde{y}|\beta; \tilde{x}, \sigma_\varepsilon^2) f(\beta|y, X, V_i, \sigma_\varepsilon^2) d\beta$$

where $f(\tilde{y}|\beta; \tilde{x}, \sigma_\varepsilon^2)$ is the pdf of $N(\tilde{x}^T\beta, \sigma_\varepsilon^2)$ and $f(\beta|y, X, V_i, \sigma_\varepsilon^2)$ is the posterior distribution of β; we denote the posterior mean and variance β_n and Σ_n, respectively.

Since both of these are normal, integrating out β leads to a normal distribution with mean

$$E[\tilde{y}|y] = E[E[\tilde{y}|\beta, y]] = E[\tilde{x}^T\beta|y] = \tilde{x}^T\beta_n$$

and with variance

$$\begin{aligned}\text{Var}[\tilde{y}|y] &= E[\text{Var}[\tilde{y}|\beta, y]] + \text{Var}[E[\tilde{y}|\beta, y]] \\ &= \sigma_\varepsilon^2 + \text{Var}[\tilde{x}^T\beta|y] = \sigma_\varepsilon^2 + \tilde{x}^T\Sigma_n\tilde{x}\end{aligned}$$

Generating Sequential M-MED with data

Consider a sequence of T steps:

1. Choose N_1 initial points using M-MED (without data), with $\alpha = 0$; call it D_1. This amounts to initializing with a space-filling design.
2. Collect data on D_1; call it y_1.
3. Choose N_2 additional points using M-MED (with data), with $\alpha = \frac{2p}{T}t$, $t = 1 : T$; call these design points D_2.
4. Collect data, y_2, on D_2.
5. Repeat steps 3-4 for $t = 2 : T$.

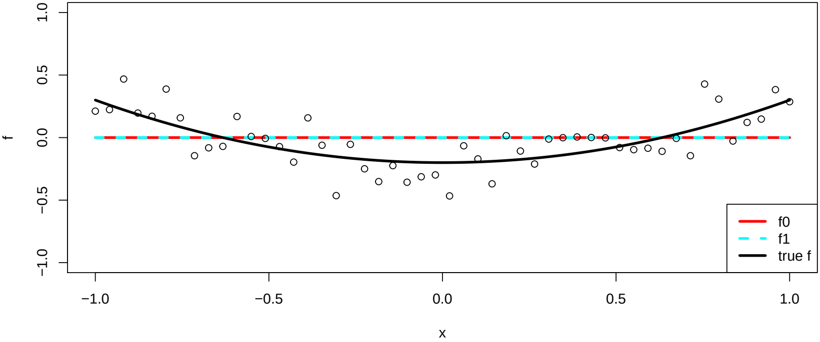
An Example

Most Basic Example

Just to see what's happening, I choose:

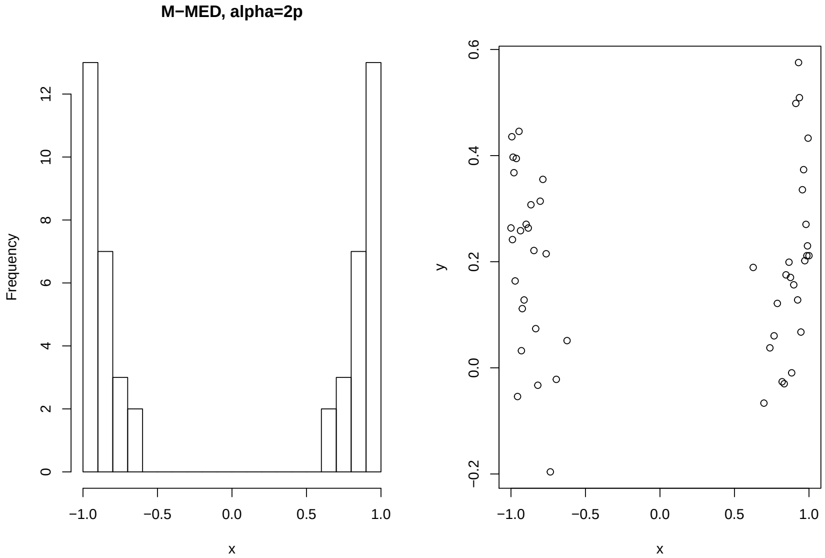
```
mu0 = c(0,0)
mu1 = c(0,0,0)
sigmasq01 = 0.01
sigmasq = 0.01

betaT = c(-0.2, 0, 0.5) # like ex2.1
```



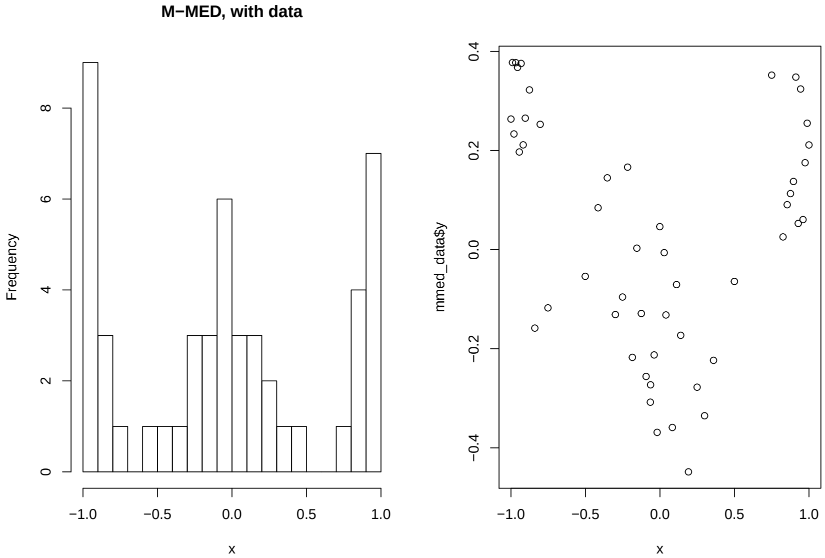
M-MED Without Data

I generate 50 design points from M-MED with no data, one with $\alpha = 2p$.

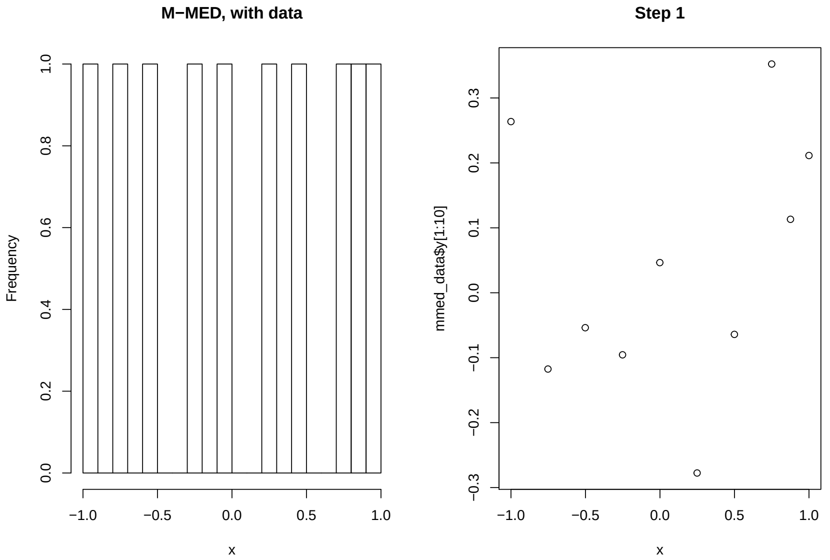


M-MED With Data

I work with a sequence of 5 steps, generating 10 points in each step, and resulting in 50 points, too.

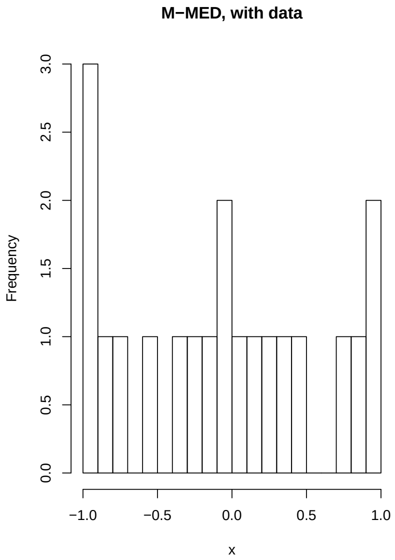


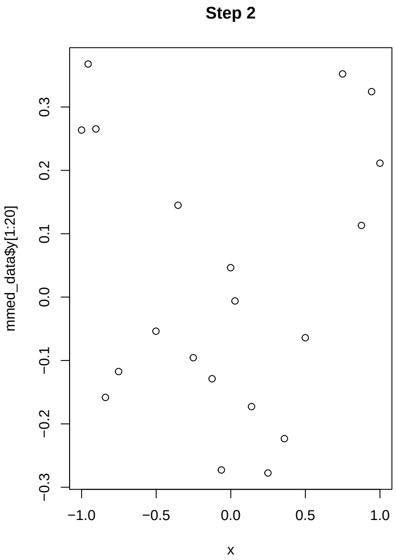
Step 1



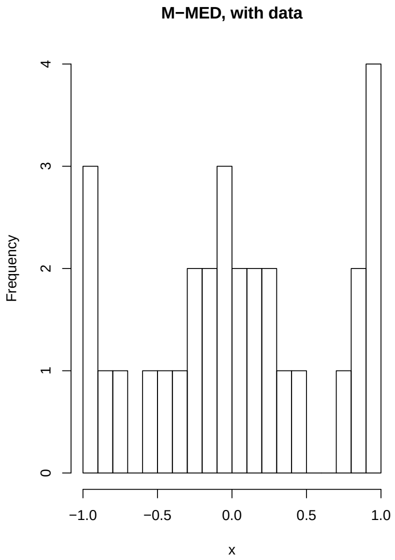
M−MED, with data
Frequency
0.0
0.2
0.4
0.6
0.8
1.0
−1.0
−0.5
0.0
0.5
1.0
x
Step 1
mmed_data$y[1:10]
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
−1.0
−0.5
0.0
0.5
1.0
x

Step 2

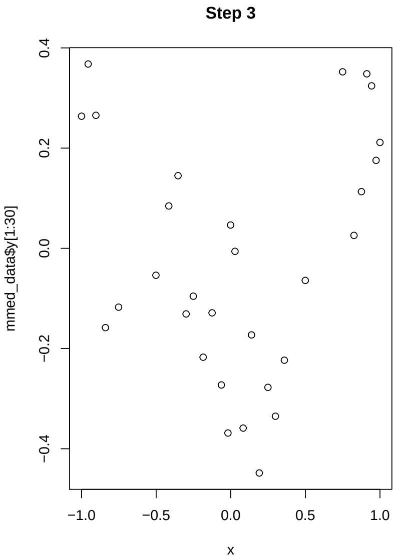

M−MED, with data
Frequency
0.0
0.5
1.0
1.5
2.0
2.5
3.0
−1.0
−0.5
0.0
0.5
1.0
x


Step 2
mmed_data$y[1:20]
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
−1.0
−0.5
0.0
0.5
1.0
x

Stage 3

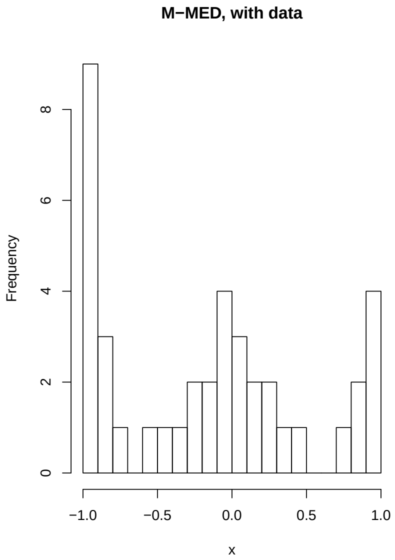


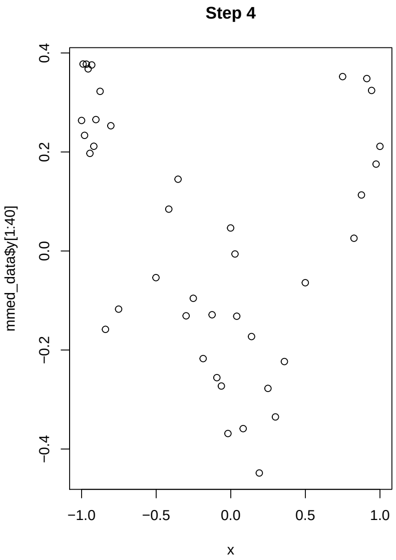
M−MED, with data
Frequency
0
1
2
3
4
−1.0
−0.5
0.0
0.5
1.0
x



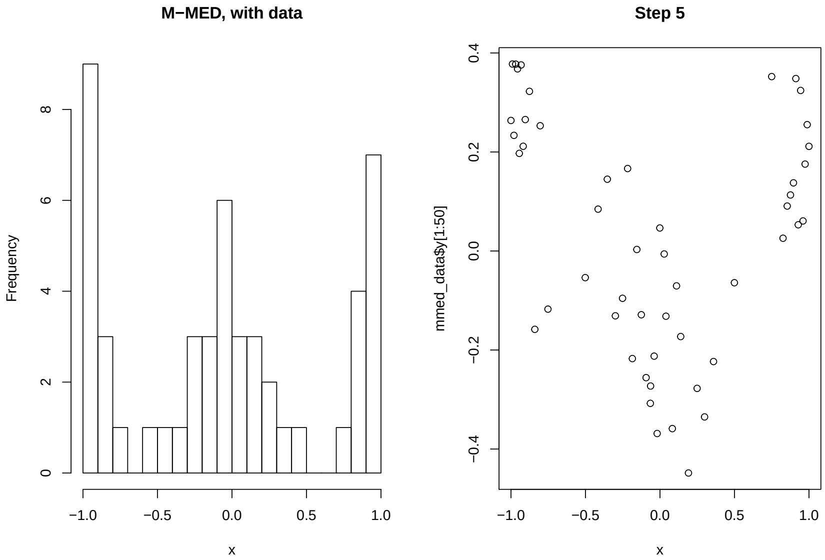
Step 3
mmed_data$y[1:30]
−0.4
−0.2
0.0
0.2
0.4
−1.0
−0.5
0.0
0.5
1.0
x

Stage 4


M−MED, with data
Frequency
0
2
4
6
8
−1.0
−0.5
0.0
0.5
1.0
x


Step 4
mmed_data$y[1:40]
−0.4
−0.2
0.0
0.2
0.4
−1.0
−0.5
0.0
0.5
1.0
x

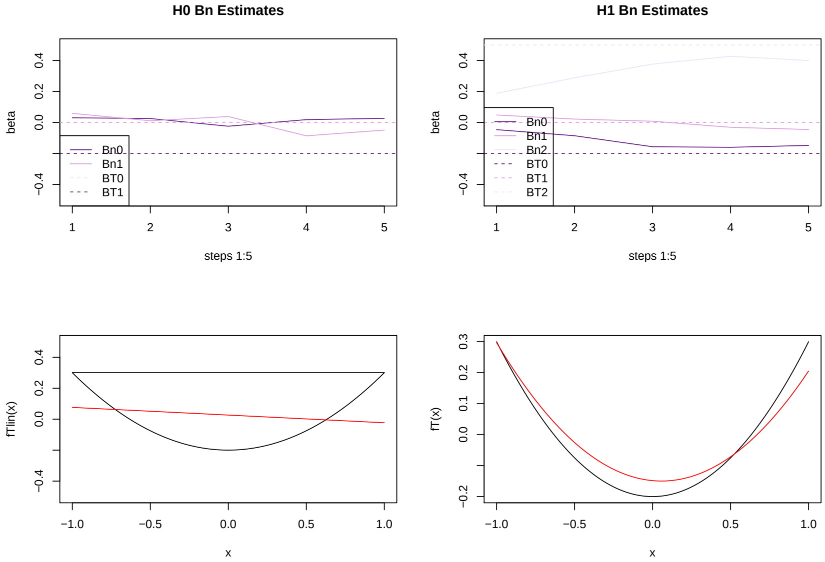
Stage 5


M−MED, with data
Frequency
0
2
4
6
8
−1.0
−0.5
0.0
0.5
1.0
x
Step 5
mmed_data$y[1:50]
−0.4
−0.2
0.0
0.2
0.4
−1.0
−0.5
0.0
0.5
1.0
x

Bn

Seeing How Bn Changes

From this one data set



MSE(Bn)

Closed Form MSE of Posterior Mean

For notation, call $E[\beta|Y] = \beta_n$.

$$\begin{aligned} MSE(\beta_n) &= Var[\beta_n] + (E[\beta_n] - \beta_T)^2 \\ &= Var[\beta_n] + (E[\beta_n])^2 - 2\beta_T E[\beta_n] + \beta_T^2 \end{aligned}$$

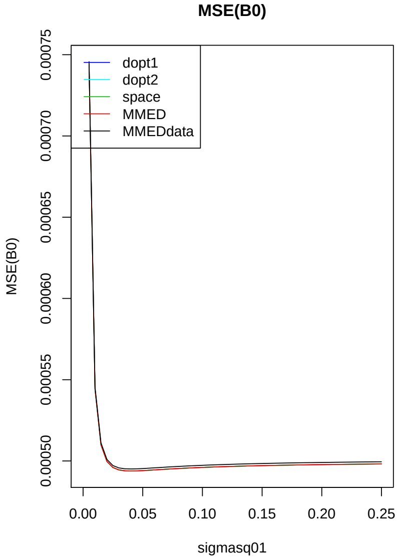
where

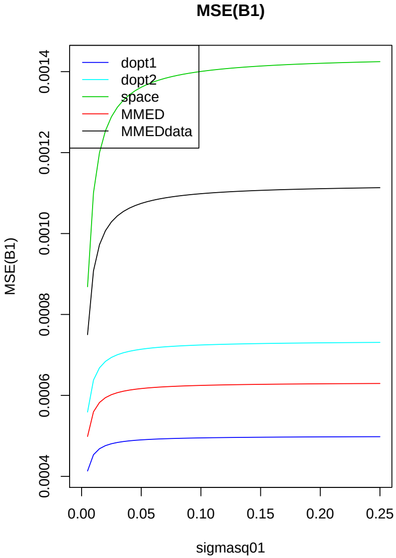
$$\begin{aligned} Var[\beta_n] &= Var[\frac{1}{\sigma^2}\Sigma_B(X^T y + \sigma^2 V^{-1}\mu)] = Var[\frac{1}{\sigma^2}\Sigma_B X^T y] \\ &= (\frac{1}{\sigma^2})^2 \Sigma_B X^T Var[y] X \Sigma_B = (\frac{1}{\sigma^2})^2 \Sigma_B X^T (\sigma^2 I) X \Sigma_B \\ &= \frac{1}{\sigma^2}\Sigma_B X^T X \Sigma_B \\ E[\beta_n] &= E[\frac{1}{\sigma^2}\Sigma_B(X^T y + \sigma^2 V^{-1}\mu)] = \frac{1}{\sigma^2}\Sigma_B(X^T E[y] + \sigma^2 V^{-1}\mu) \\ &= \frac{1}{\sigma^2}\Sigma_B(X^T X\beta_T + \sigma^2 V^{-1}\mu) = \frac{1}{\sigma^2}\Sigma_B X^T X \beta_T + \Sigma_B V^{-1}\mu \end{aligned}$$

where $\Sigma_B = Var[\beta|Y] = \sigma^2(X^T X + \sigma^2 V^{-1})^{-1}$ and
$y \sim N(X\beta_T, \sigma^2 I)$

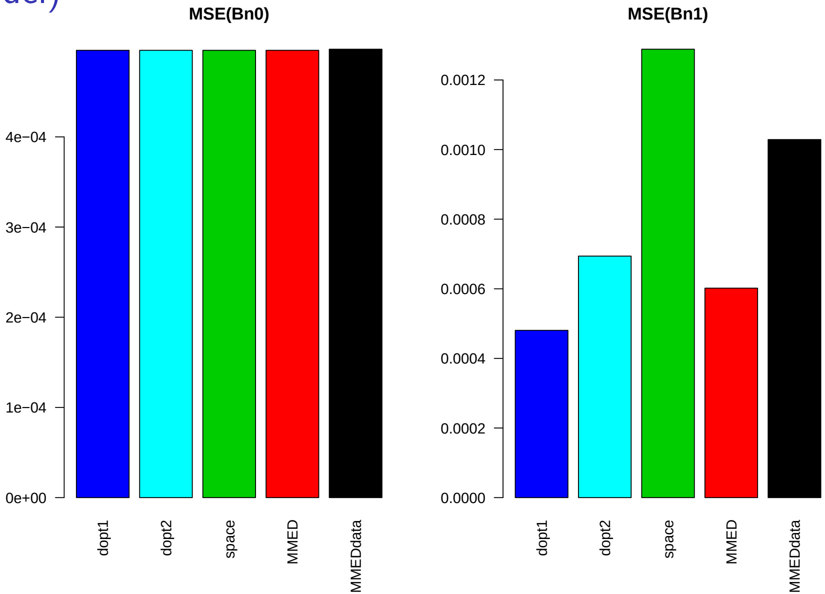
MSE(Bn) (assuming linear model)

This doesn't seem to match what we see for B_n and $E[B_n]$ (which, come to think, can be calculated in closed form too)

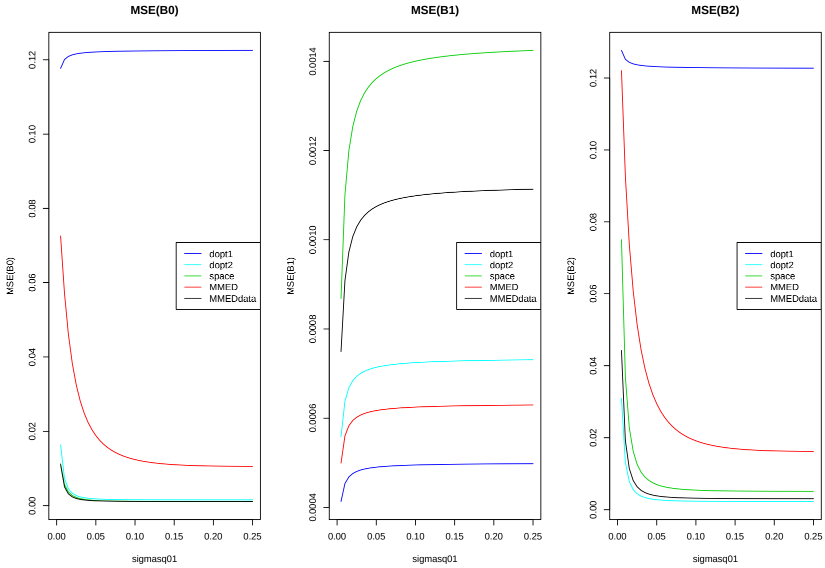




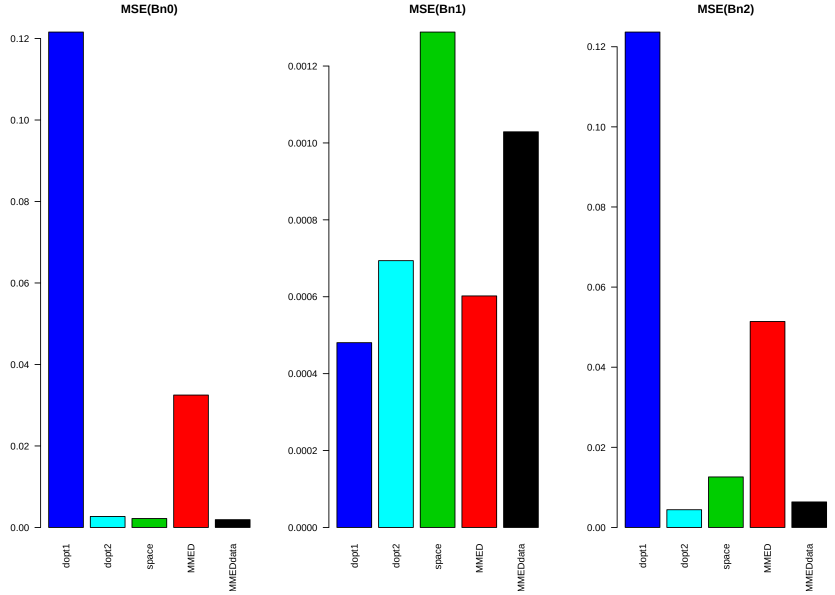
MSE Bar Plots (since sigma01 is fixed) (assuming linear model)



MSE(Bn) (assuming quadratic model)



MSE Bar Plots (assuming quadratic model)

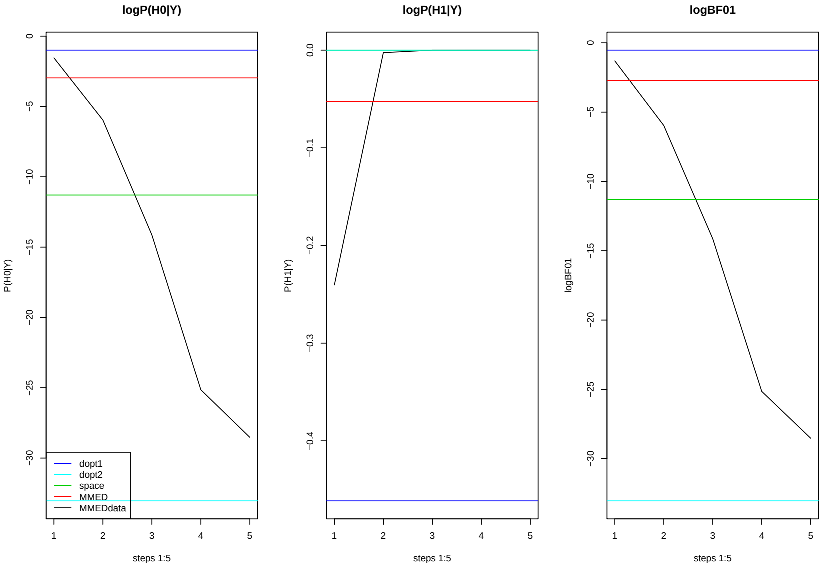


Bn and EBn

(under construction)

Posterior Probabilities of Hypotheses

Hypothesis Testing



```r
exppostprobs_dopt1
```

```
## expected_postH0 expected_postH1   expected_BF01
##       0.3696449       0.6303551       0.5896454
```

```r
exppostprobs_dopt2
```

```
## expected_postH0 expected_postH1   expected_BF01
##    4.469577e-15    1.000000e+00    4.469577e-15
```

```r
exppostprobs_mmednodata
```

```
## expected_postH0 expected_postH1   expected_BF01
##      0.05138706      0.94861294      0.06519469
```

```r
exppostprobs_space
```

```
## expected_postH0 expected_postH1   expected_BF01
##    1.241255e-05    9.999876e-01    1.241748e-05
```

```r
c(postprobs0[5], postprobs1[5], BF01s[5])
```

```
## [1] 4.085207e-13 1.000000e+00 4.085207e-13
```