

Evaluations for Examples

Kristyn Pantoja

9/29/2019

1 Dimension

Example 1

Example 2

Example 3

Example 4

Example 4.1

Example 5

Summary

- ▶ In these slides, I look at different pairs of sharp hypotheses for generating MED, and how they evaluate based on
 - ▶ $E[P(H_1|Y, D)|\beta_T]$
 - ▶ $MSE(\beta_n|Y, D, H_1, \beta_T)$
- ▶ And we try to find an example where the quadratic D-Optimal design isn't best for the $E[P(H_1|Y, D)|\beta_T]$ evaluations. (Example 4.1, but also 1 & 5)
- ▶ This time, $\beta_T \sim N(\mu_1, V_1)$ (whereas last time, $\beta_T = \mu_1$)
- ▶ Things to consider:
 - ▶ For posterior evaluations, when $\beta_T = \mu_T$, it might not make sense to compare the D-optimal designs
 - ▶ We want to pay attention to cases of $\beta_T \neq \mu_T$ where D-optimal designs are expected to do worse than MEDs
 - ▶ And find cases where space-filling design does worse than MEDs

Observations from Last Time (Example 1: Estimating Linear Term's Coefficient)

In last slides, we saw what happens when $\beta_T = \mu_T$.

- ▶ In Example 1, where the MED concentrates at -1 and 1, we saw that:
 - ▶ for the linear term's coefficient, $MSE(\beta_{n,1}|Y, D, H_1, \beta_T)$ was smaller (i.e. better) for the MED than for space-filling design, but
 - ▶ for the intercept and quadratic term's coefficients, $MSE(\beta_{n,0}|Y, D, H_1, \beta_T)$ and $MSE(\beta_{n,2}|Y, D, H_1, \beta_T)$ were larger (i.e. worse) for MED compared to the space-filling design except at smaller values of the prior variance on β .
 - ▶ This makes me think that, like the linear D-optimal design suggests, concentrating at -1 and 1 is best for estimating the linear term - however, it is not best for estimating the intercept and quadratic term.

Observations from Last Time (Example 2: Estimating Intercept and Quadratic Term's Coefficient)

- ▶ In Example 2, where MED concentrates at -1, 1, and 0 (but mostly at 0), we saw that:
 - ▶ For $E[P(H_1|Y, D)|\beta_T]$, MED seems to be a compromise between the quadratic D-optimal design and the space-filling design (yay!)
 - ▶ For $MSE(\beta_n|Y, D, H_1, \beta_T)$,
 - ▶ $MSE(\beta_{n,0}|Y, D, H_1, \beta_T)$: MED is better than space-filling design
 - ▶ $MSE(\beta_{n,1}|Y, D, H_1, \beta_T)$: MED, surprisingly, is worse than space-filling design (perhaps it focuses too much in the middle than at the edges)
 - ▶ $MSE(\beta_{n,2}|Y, D, H_1, \beta_T)$: MED and space-filling are comparable. . . MED with $\alpha = 1$ does better than space-filling design, but MED with $\alpha = 2p$ does worse than space-filling design

Observations from Last Time (Example 3: Estimating Intercept well)

- ▶ In Example 3, MEDs concentrate at 0
 - ▶ For $E[P(H_1|Y, D)|\beta_T]$, the space-filling design does worse than the MED with $\alpha = 1$, but a lot better than the MED $\alpha = 2\rho$, which was disturbing.
 - ▶ For $MSE(\beta_n|Y, D, H_1, \beta_T)$,
 - ▶ $MSE(\beta_{n,0}|Y, D, H_1, \beta_T)$: As expected, given what we saw in Example 2, MEDs did better than the space-filling design, since they concentrate more points at 0
 - ▶ For $MSE(\beta_{n,1}|Y, D, H_1, \beta_T)$ and $MSE(\beta_{n,2}|Y, D, H_1, \beta_T)$, though, MED did worse than the space-filling design.

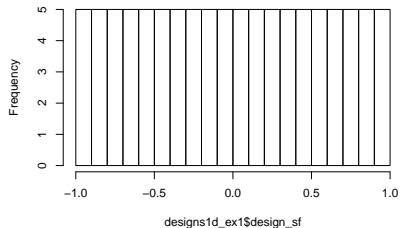
Why look at different BetaT

- ▶ $MSE(\beta_n | Y, D, H_1, \beta_T)$ evaluations: we want to show examples where the D-optimal design doesn't do well. Since D-optimal designs specify a general model, it seems that it would require specific models to show that it's not "optimal" in certain senses.

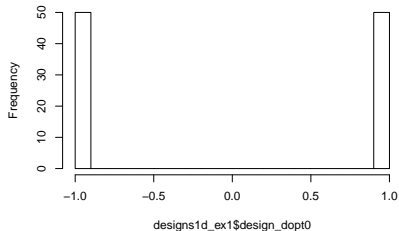
1 Dimension

Non-MED Designs

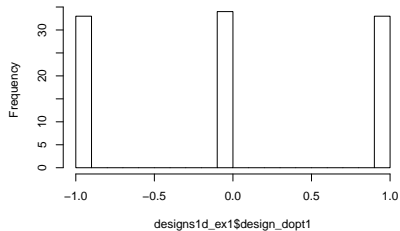
Space-Filling



D-Optimal, Linear



D-Optimal, Quadratic



Example 1

Example 1 Parameters

- ▶ The idea here is a model with no assumptions about the value for the quadratic term. The result is a more spread-out version of the linear D-optimal design.
- ▶ Idea: Should we instead apply this to model comparison with no assumptions about the value for the linear term? i.e. $\mu_0 = (0)$ and $\mu_1 = (0, 0)$?
- ▶ In any case...

Priors

```
sigmasq01 = 0.01
```

```
mu0 = c(0,1)
```

```
V0 = diag(rep(sigmasq01,length(mu0)))
```

```
mu1 = c(0,1,0)
```

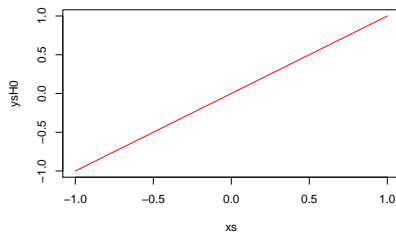
```
V1 = diag(rep(sigmasq01,length(mu1)))
```

```
sigmasq = 1
```

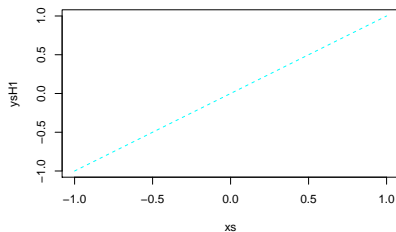
```
betaT = c(-5, 1, 5)
```

Example 1 Plots

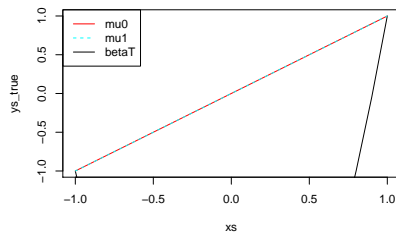
H0 model



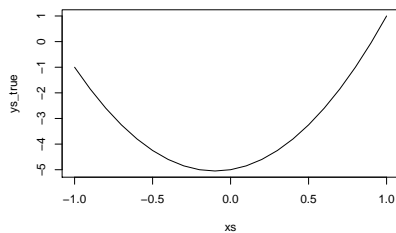
H1 model



true model

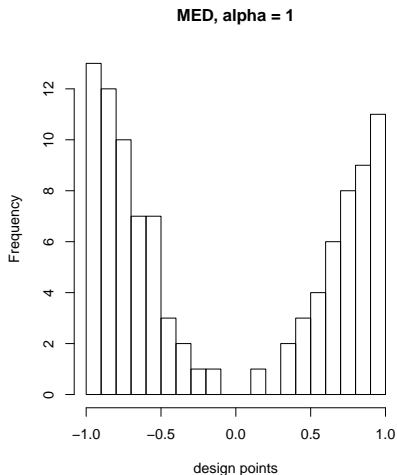
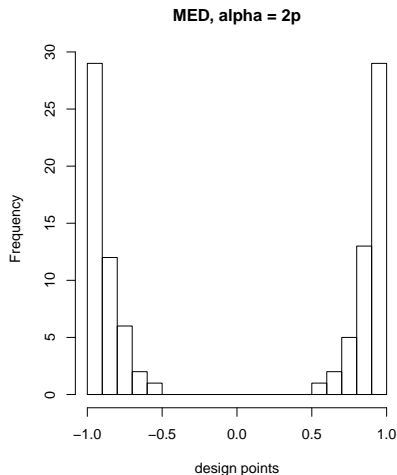


true model, full view



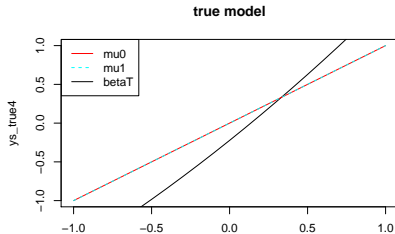
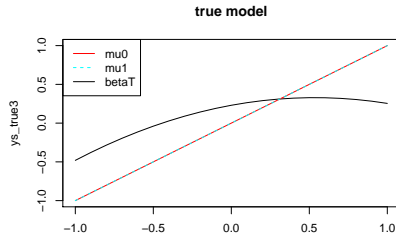
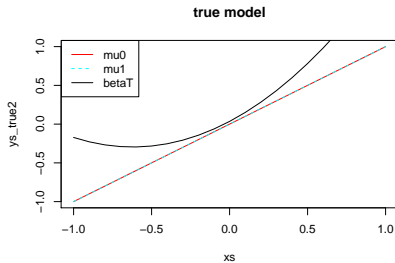
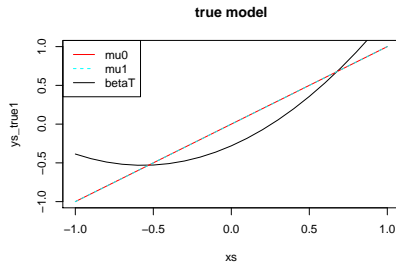
Example 1 MEDs

- ▶ The original MED (with $\alpha = 1$) has a more space-filling tendency than the adjusted MED (with no power on $f_D(\cdot, \cdot) = \text{Wasserstein}(\cdot, \cdot)$, i.e. $\alpha = 2p$).



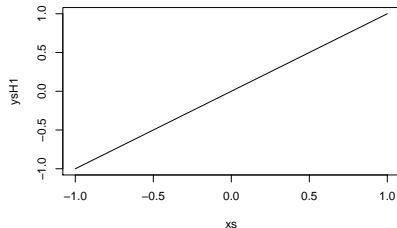
Example 1, Different BetaT

Now we generate different values of $\beta_T \sim N((0, 1, 0)^T, V_1)$ using $\text{sigmasq01} = 0.25$. (should sigmasq01 be larger for more variety?)

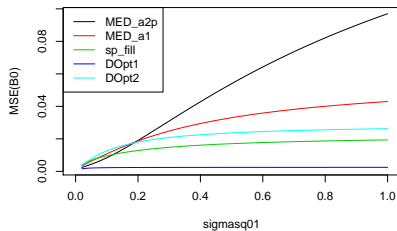


MSEs, $BT = \mu_1$

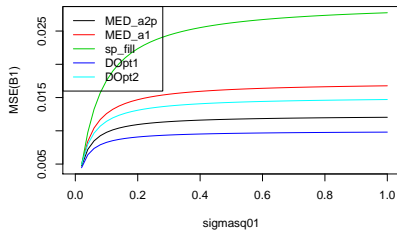
true model = μ_1



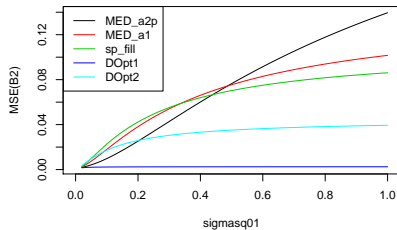
$BT=(0,1,0)$, $MSE(B0)$



$BT=(0,1,0)$, $MSE(B1)$



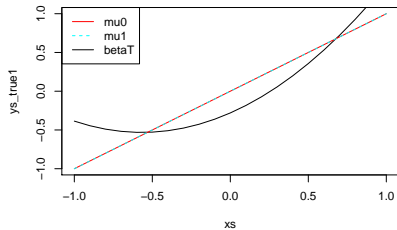
$BT=(0,1,0)$, $MSE(B2)$



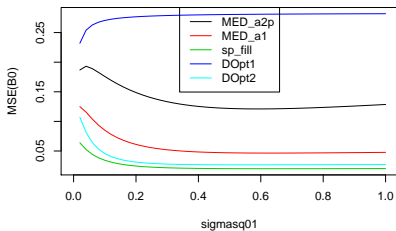
MSEs, BT(1)

betaTs[,1]

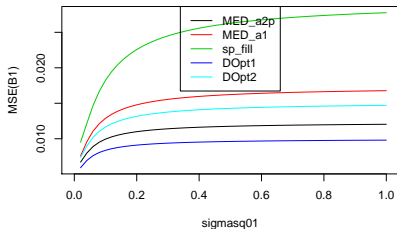
true model



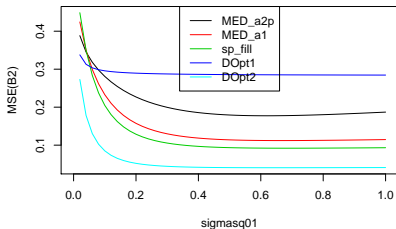
BT1=(-0.28,0.885,0.779), MSE(B0)



BT1=(-0.28,0.885,0.779), MSE(B1)



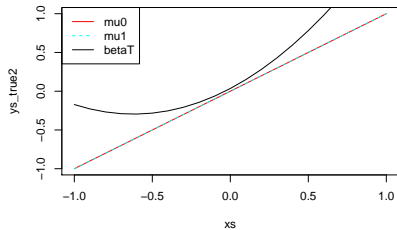
BT1=(-0.28,0.885,0.779), MSE(B2)



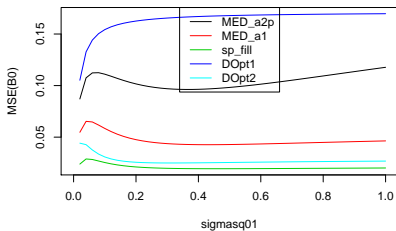
MSEs, BT(2)

betaTs[,2]

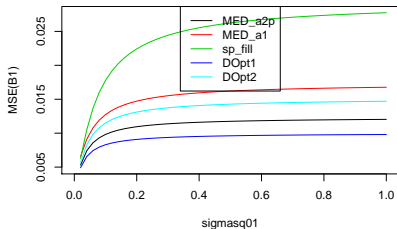
true model



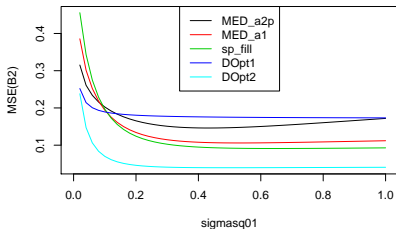
BT1=(0.035,1.065,0.858), MSE(B0)



BT1=(0.035,1.065,0.858), MSE(B1)



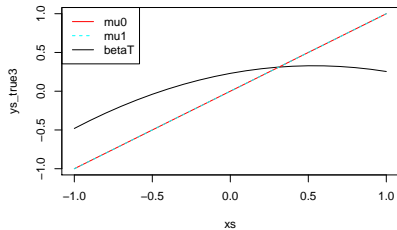
BT1=(0.035,1.065,0.858), MSE(B2)



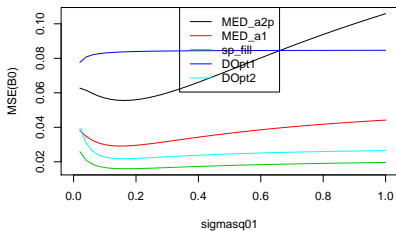
MSEs, BT(3)

betaTs[,3]

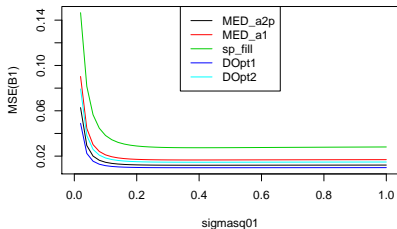
true model



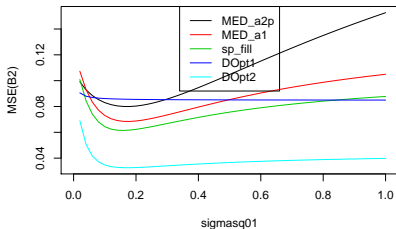
BT1=(0.23,0.367,-0.343), MSE(B0)



BT1=(0.23,0.367,-0.343), MSE(B1)

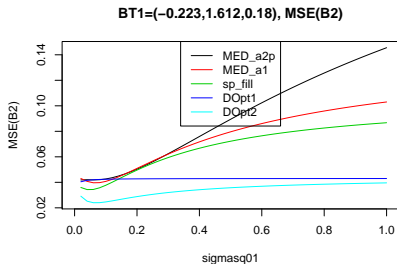
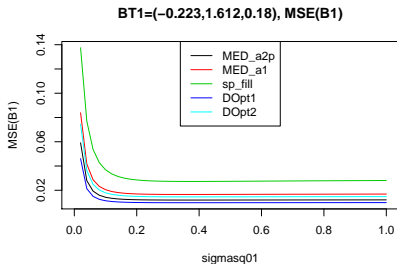
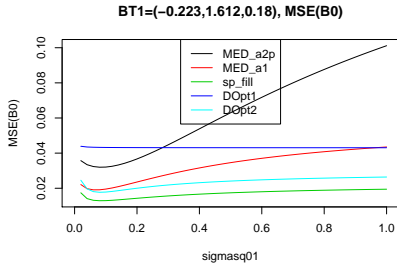
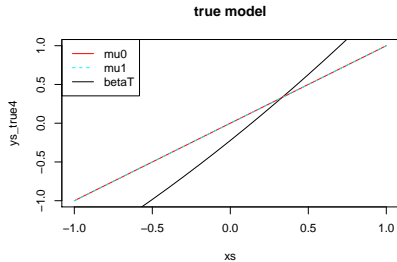


BT1=(0.23,0.367,-0.343), MSE(B2)

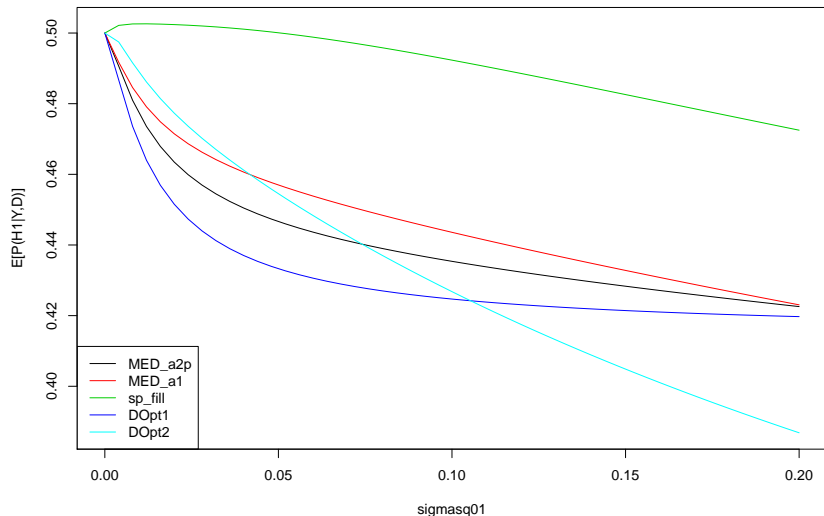


MSEs, BT(4)

betaTs[,4]

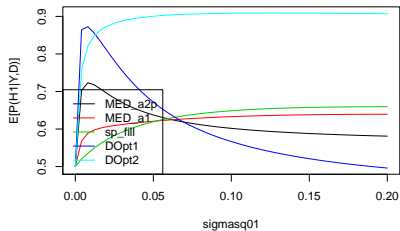


Example 1 $E[P(H1|Y,D, \text{beta}T=\mu1)]$

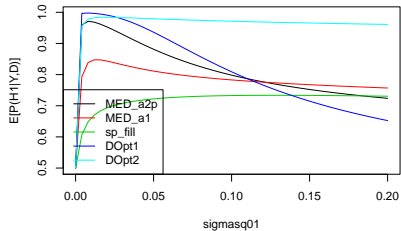


Example 1 $E[P(H1|Y,D, \text{beta}T)]$

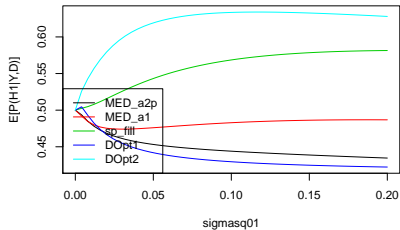
BT1=(-0.28,0.885,0.779)



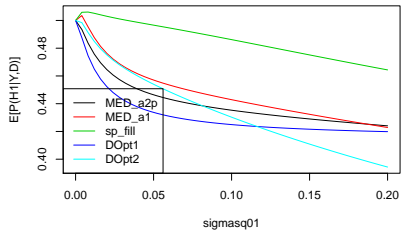
BT2=(0.035,1.065,0.858)



BT3=(0.23,0.367,-0.343)



BT4=(-0.223,1.612,0.18)



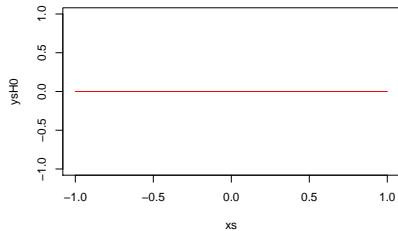
Example 2

Example 2 Parameters

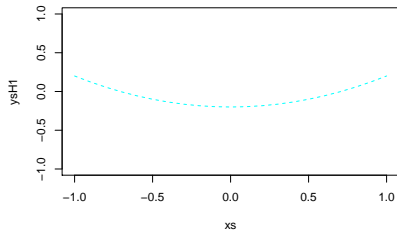
```
# Priors
sigmasq01 = 0.01
mu0 = c(0, 0)
V0 = diag(rep(sigmasq01,length(mu0)))
mu1 = c(-0.2, 0, 0.4)
V1 = diag(rep(sigmasq01,length(mu1)))
sigmasq = 1
betaT = mu1
```

Example 2 Plots

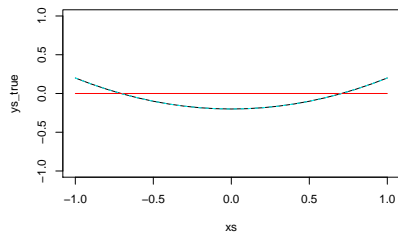
H0 model



H1 model

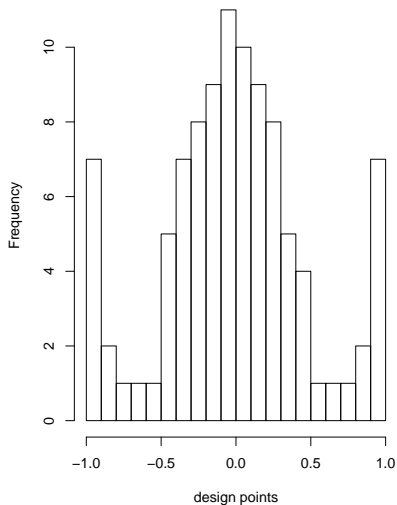


true model

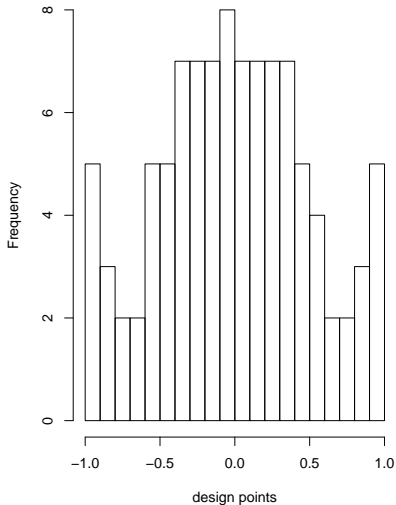


Example 2 MEDs

MED, $\alpha = 2p$

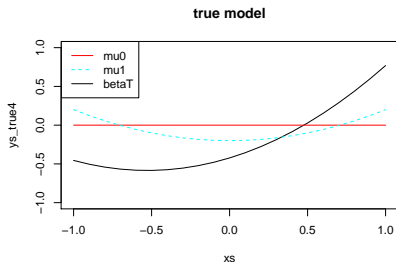
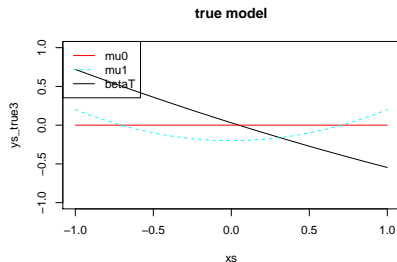
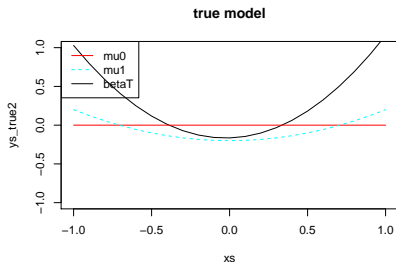
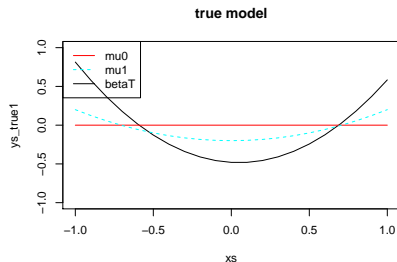


MED, $\alpha = 1$



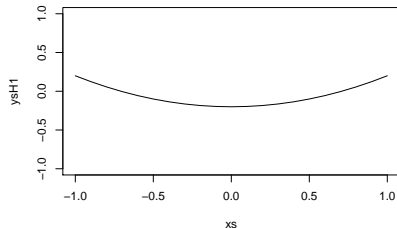
Example 2, Different BetaT

Now we generate different values of $\beta_T \sim N((-0.2, 0, 0.4)^T, V_1)$ using $\text{sigmasq01} = 0.1$

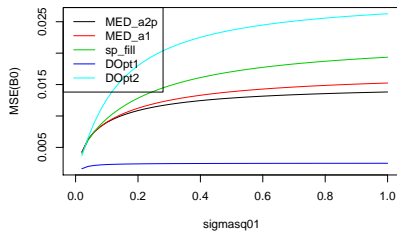


MSEs, $BT = \mu_1$

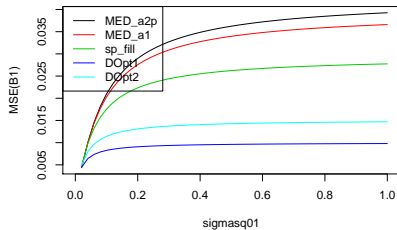
true model = μ_1



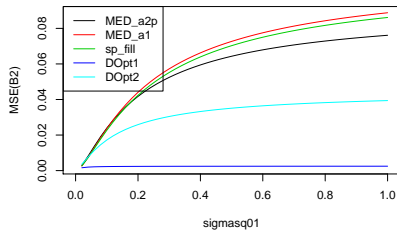
$BT = (-0.2, 0, 0.4)$, $MSE(B_0)$



$BT = (-0.2, 0, 0.4)$, $MSE(B_1)$



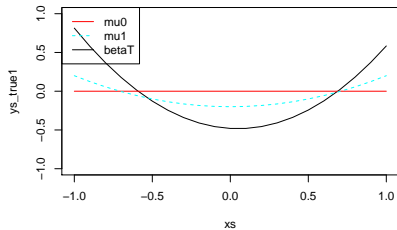
$BT = (-0.2, 0, 0.4)$, $MSE(B_2)$



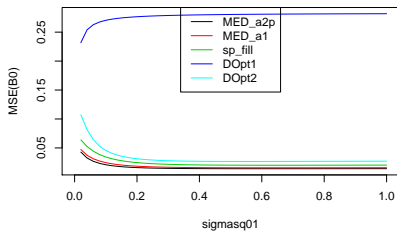
MSEs, BT(1)

betaTs[,1]

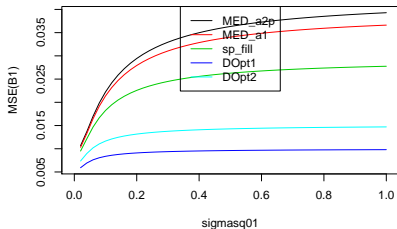
true model



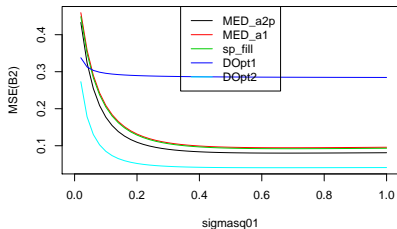
BT1=(-0.48,-0.115,1.179), MSE(B0)



BT1=(-0.48,-0.115,1.179), MSE(B1)



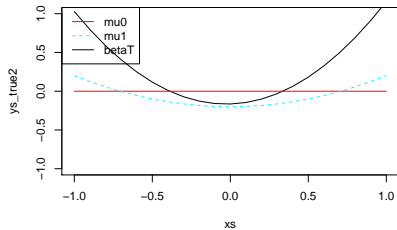
BT1=(-0.48,-0.115,1.179), MSE(B2)



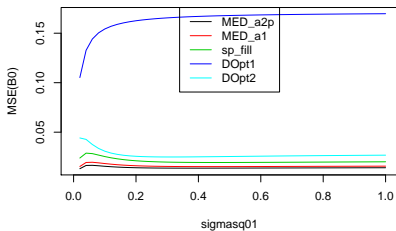
MSEs, BT(2)

betaTs[,2]

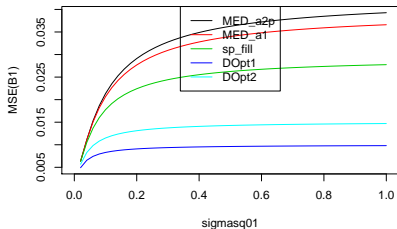
true model



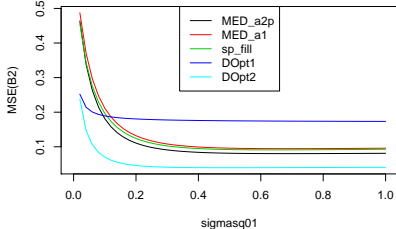
BT1=(-0.165,0.065,1.258), MSE(B0)



BT1=(-0.165,0.065,1.258), MSE(B1)



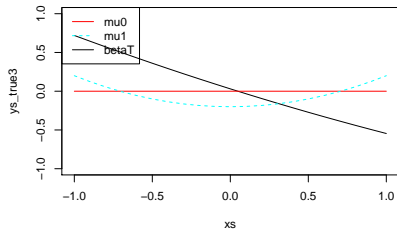
BT1=(-0.165,0.065,1.258), MSE(B2)



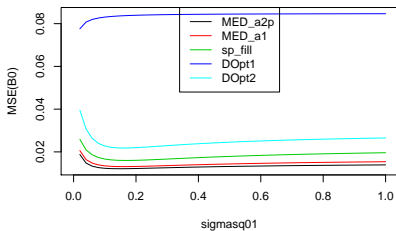
MSEs, BT(3)

betaTs[,3]

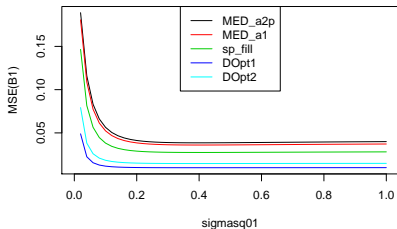
true model



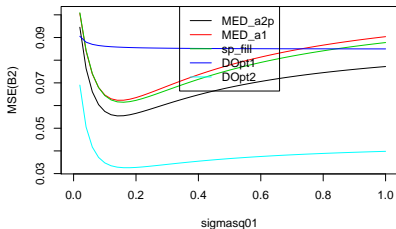
BT1=(0.03,-0.633,0.057), MSE(B0)



BT1=(0.03,-0.633,0.057), MSE(B1)



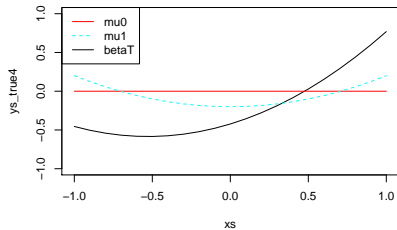
BT1=(0.03,-0.633,0.057), MSE(B2)



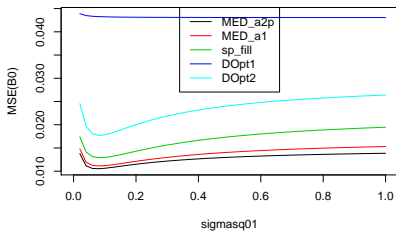
MSEs, BT(4)

betaTs[,4]

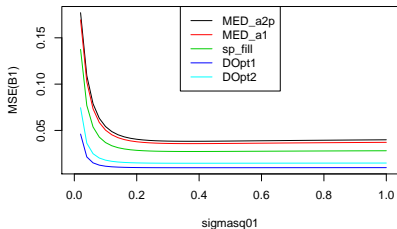
true model



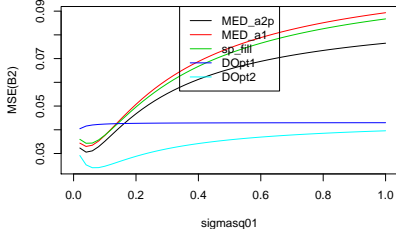
BT1=(-0.423,0.612,0.58), MSE(B0)



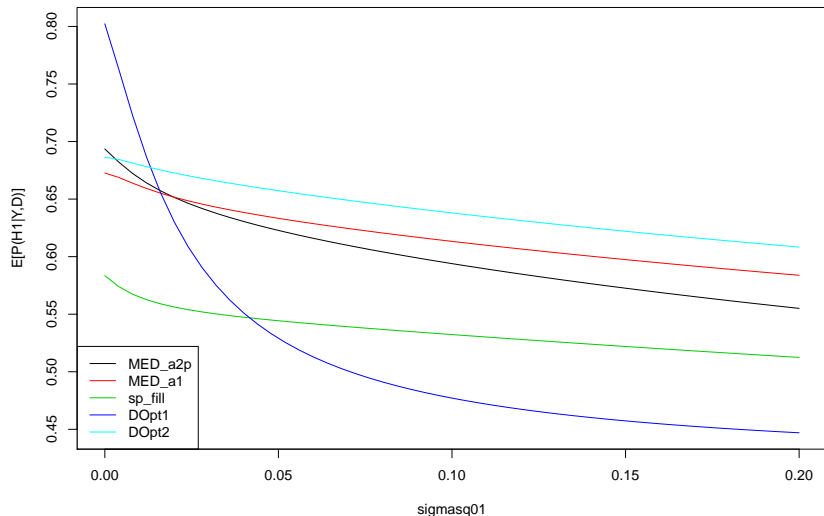
BT1=(-0.423,0.612,0.58), MSE(B1)



BT1=(-0.423,0.612,0.58), MSE(B2)

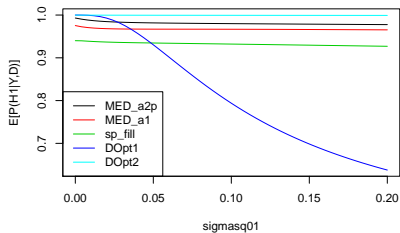


Example 2 $E[P(H1|Y,D, \text{beta}T=\mu1)]$

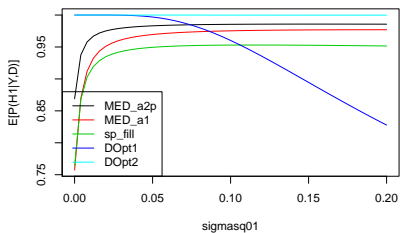


Example 2 $E[P(H1|Y,D)]$

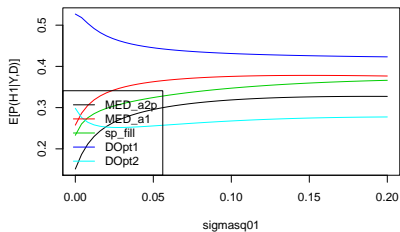
BT1=(-0.48,-0.115,1.179)



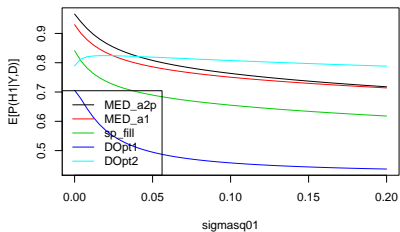
BT2=(-0.165,0.065,1.258)



BT3=(0.03,-0.633,0.057)



BT4=(-0.423,0.612,0.58)



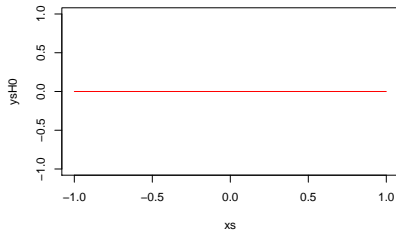
Example 3

Example 3 Parameters

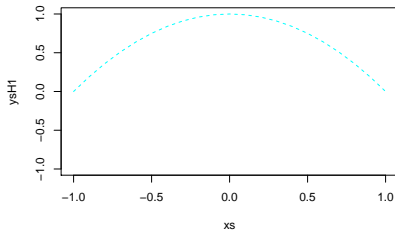
```
# Priors
sigmasq01 = 0.01
mu0 = c(0, 0, 0)
V0 = diag(rep(sigmasq01,length(mu0)))
mu1 = c(2, 0, -1, 0, -1)
V1 = diag(rep(sigmasq01,length(mu1)))
sigmasq = 1
betaT = mu1
```

Example 3 Plots

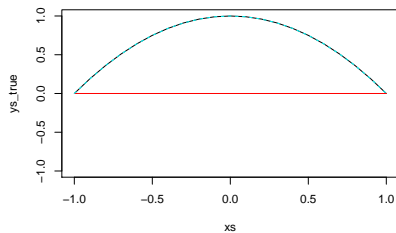
H0 model



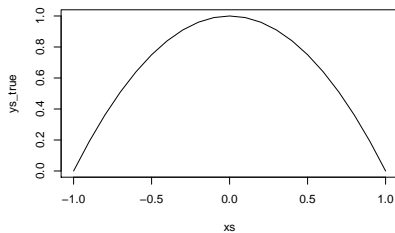
H1 model



true model

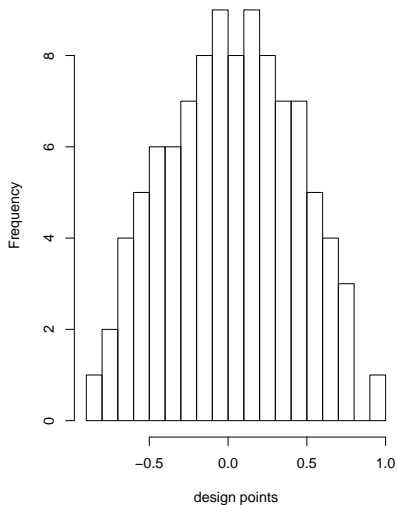


true model, full view

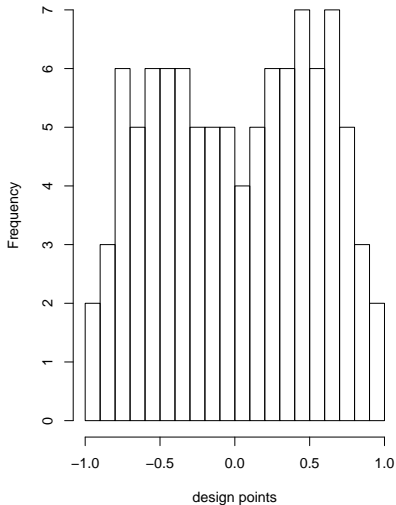


Example 3 MEDs

MED, $\alpha = 2p$

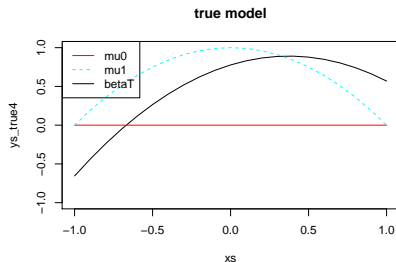
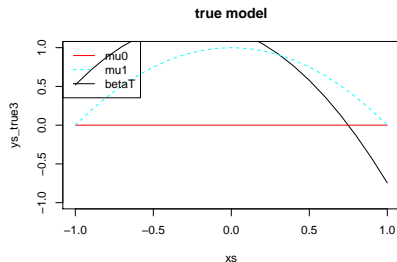
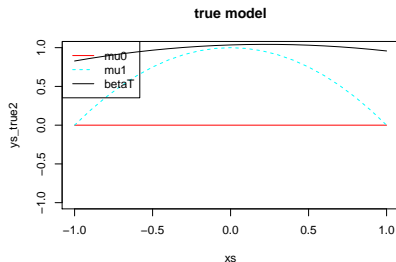
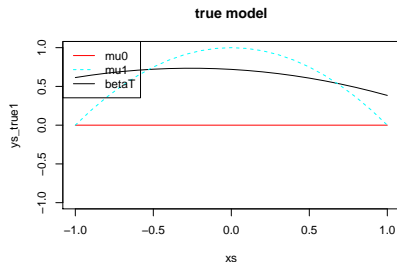


MED, $\alpha = 1$



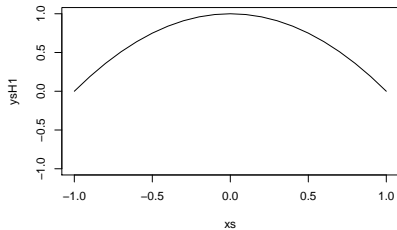
Example 3, Different BetaT

Now we generate different values of $\beta_T \sim N((0, 1, 0)^T, V_1)$ using $\text{sigmasq01} = 0.1$

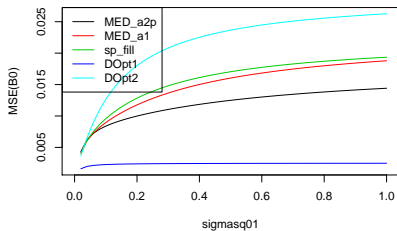


MSEs, $BT = \mu_1$

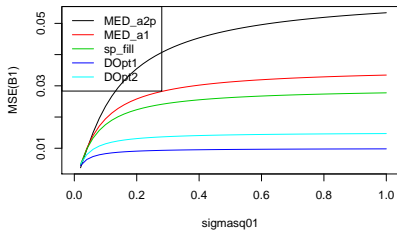
true model = μ_1



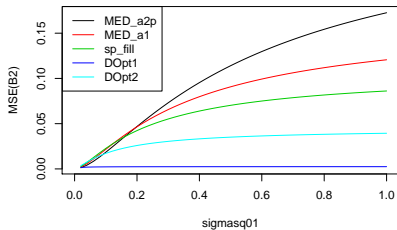
$BT=(1,0,-1)$, $MSE(B_0)$



$BT=(1,0,-1)$, $MSE(B_1)$



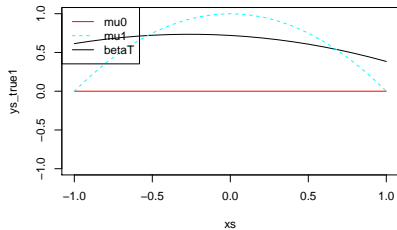
$BT=(1,0,-1)$, $MSE(B_2)$



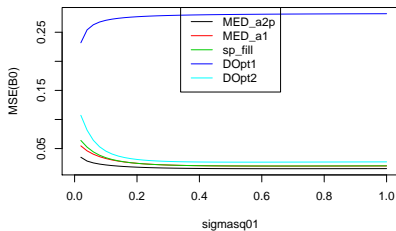
MSEs, BT(1)

betaTs[,1]

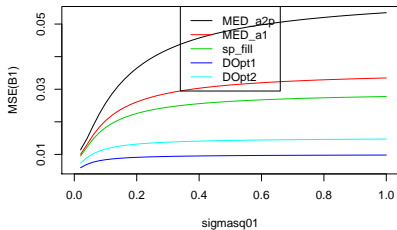
true model



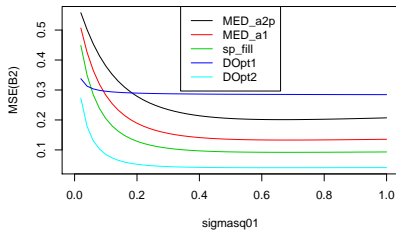
BT1=(0.72,-0.115,-0.221), MSE(B0)



BT1=(0.72,-0.115,-0.221), MSE(B1)



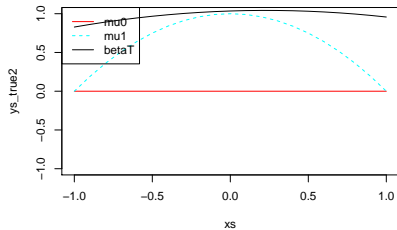
BT1=(0.72,-0.115,-0.221), MSE(B2)



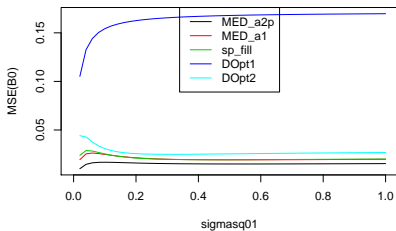
MSEs, BT(2)

betaTs[,2]

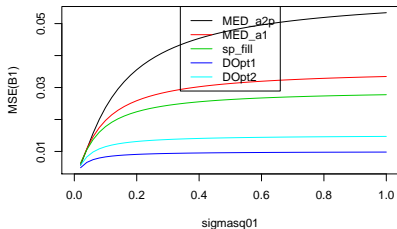
true model



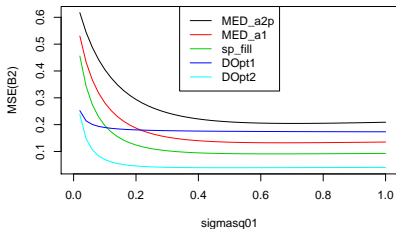
BT1=(1.035,0.065,-0.142), MSE(B0)



BT1=(1.035,0.065,-0.142), MSE(B1)

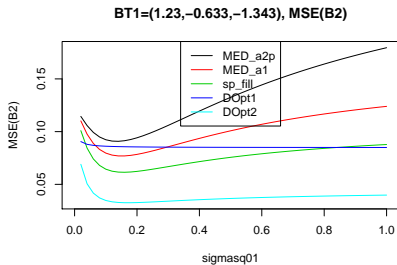
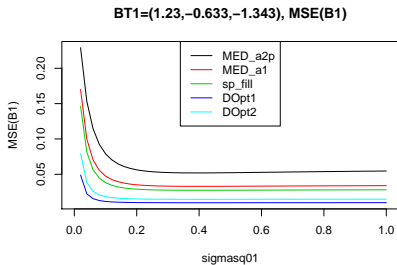
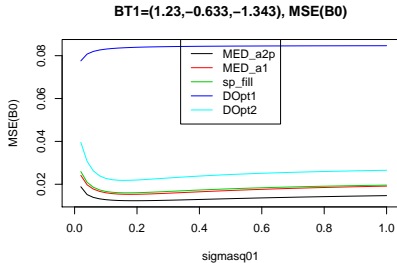
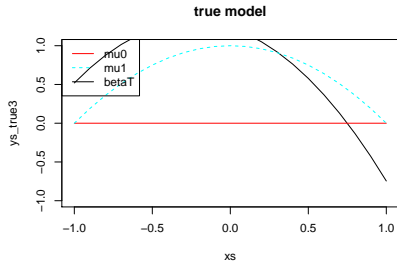


BT1=(1.035,0.065,-0.142), MSE(B2)



MSEs, BT(3)

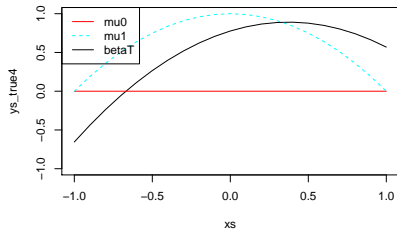
betaTs[,3]



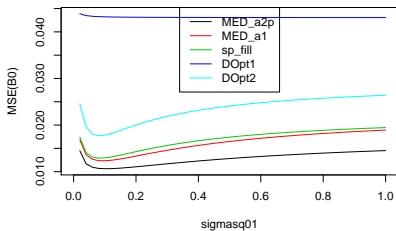
MSEs, BT(4)

betaTs[,4]

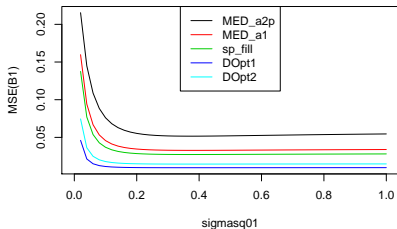
true model



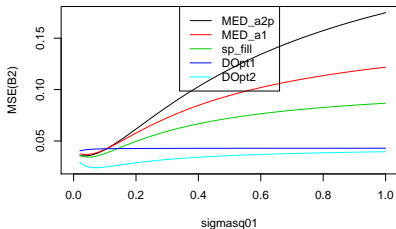
BT1=(0.777,0.612,-0.82), MSE(B0)



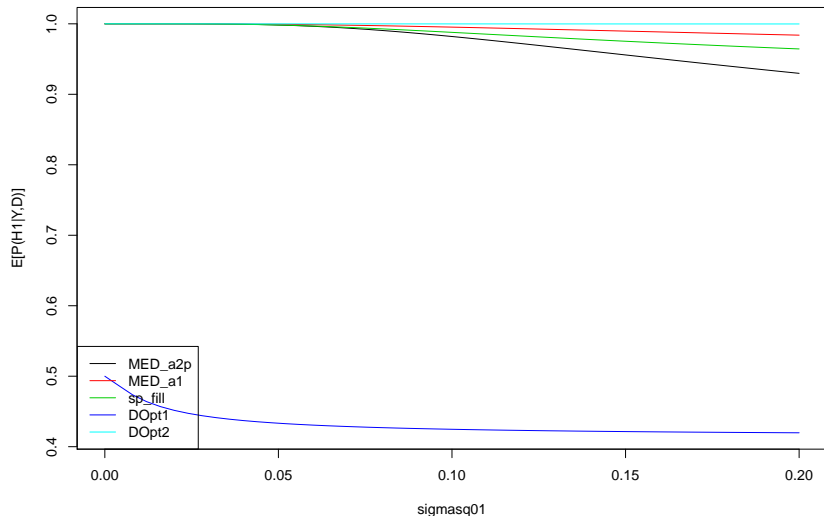
BT1=(0.777,0.612,-0.82), MSE(B1)



BT1=(0.777,0.612,-0.82), MSE(B2)

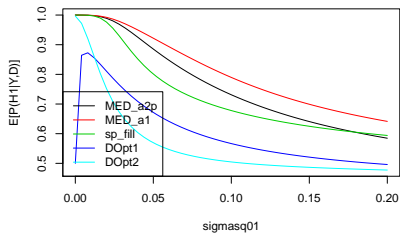


Example 3 $E[P(H1|Y,D, \text{betaT}=\mu1)]$

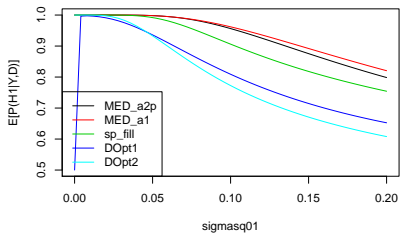


Example 3 $E[P(H1|Y,D)]$

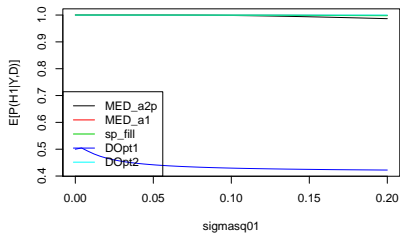
BT1=(0.72,-0.115,-0.221)



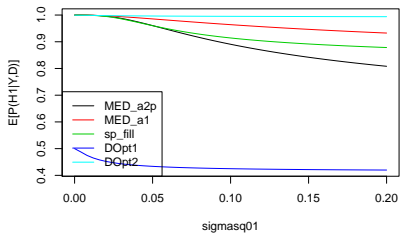
BT2=(1.035,0.065,-0.142)



BT3=(1.23,-0.633,-1.343)



BT4=(0.777,0.612,-0.82)



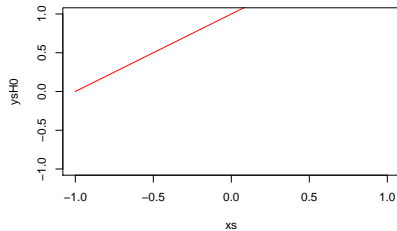
Example 4

Example 4 Parameters

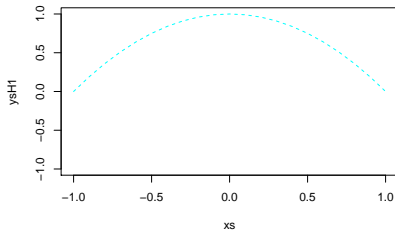
```
# Priors
sigmasq01 = 0.01
mu0 = c(1, 1)
V0 = diag(rep(sigmasq01,length(mu0)))
mu1 = c(1, 0, -1)
V1 = diag(rep(sigmasq01,length(mu1)))
sigmasq = 1
betaT = mu1
```

Example 4 Plots

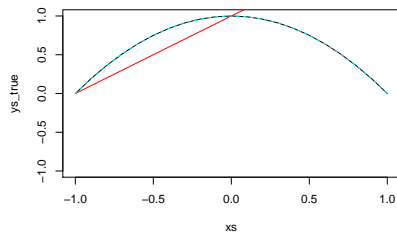
H0 model



H1 model

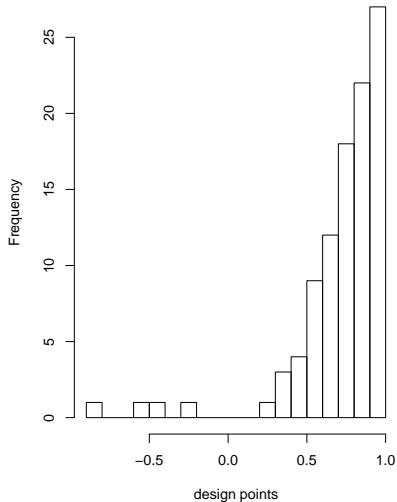


true model

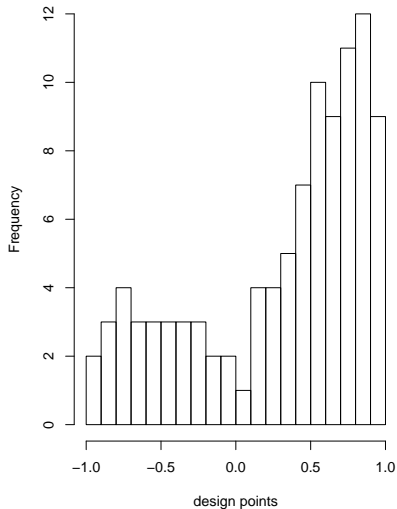


Example 4 MEDs

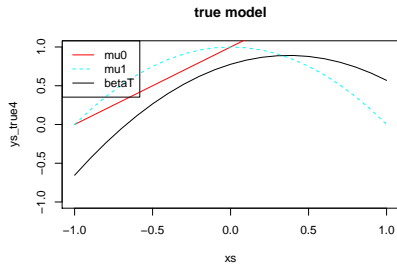
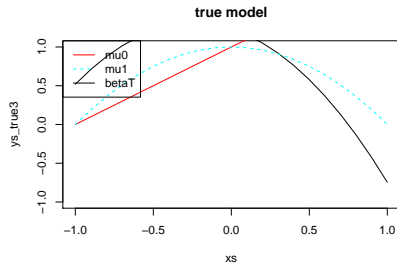
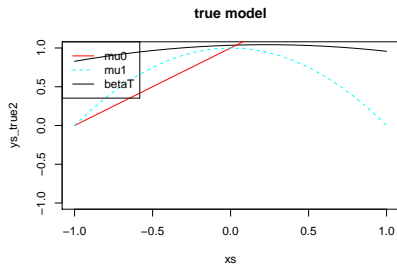
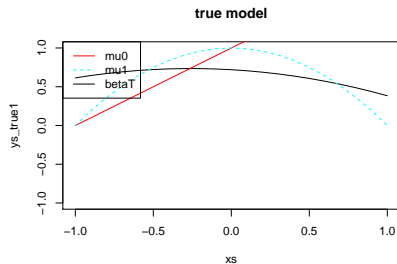
MED, $\alpha = 2p$



MED, $\alpha = 1$

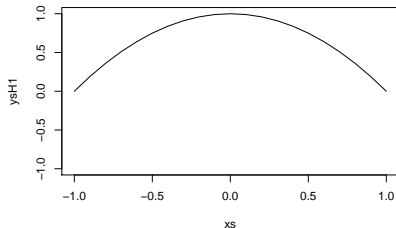


Example 4, Different BetaT

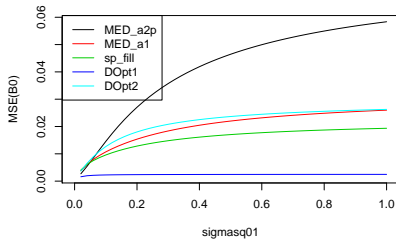


MSEs, BT = μ_1

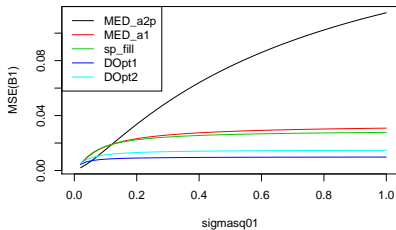
true model = μ_1



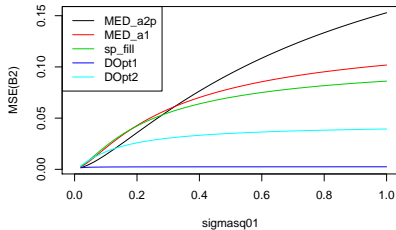
BT=(1,0,-1), MSE(B0)



BT=(1,0,-1), MSE(B1)

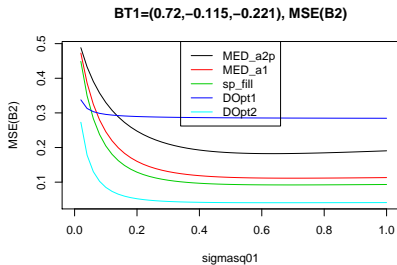
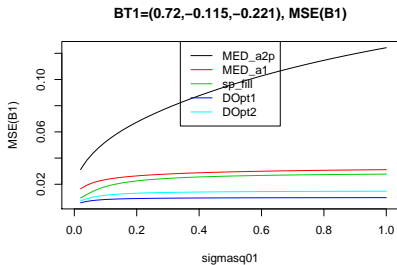
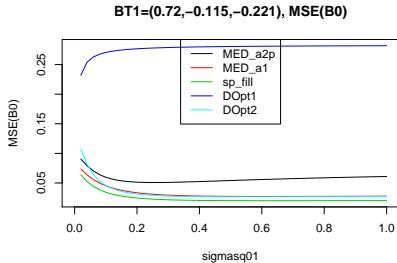
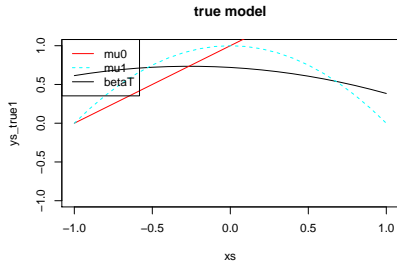


BT=(1,0,-1), MSE(B2)



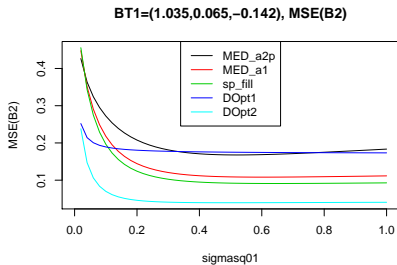
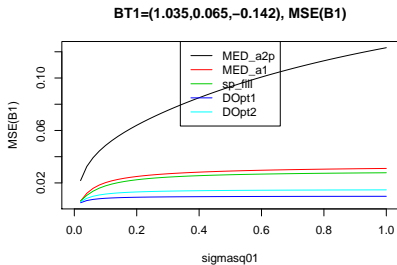
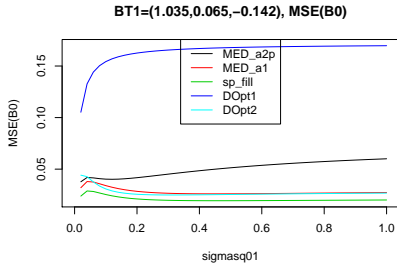
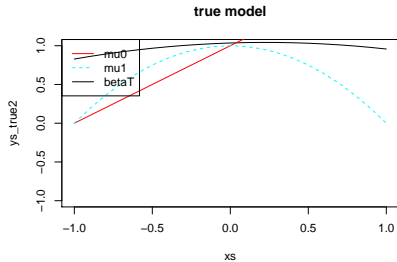
MSEs, BT(1)

betaTs[,1]



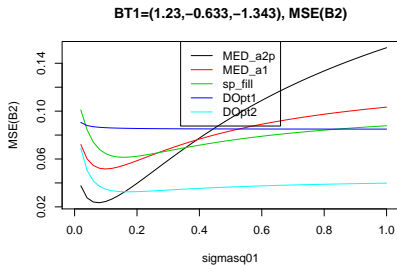
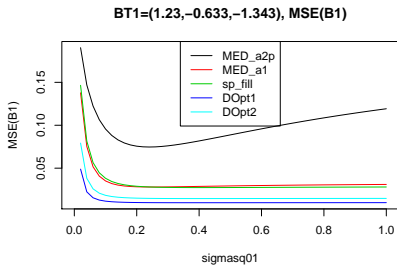
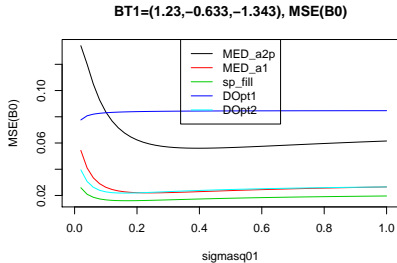
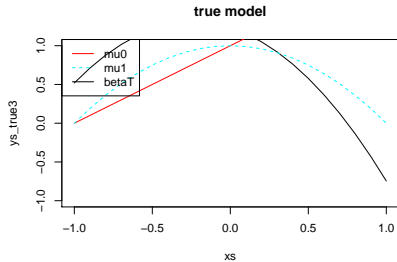
MSEs, BT(2)

betaTs[,2]



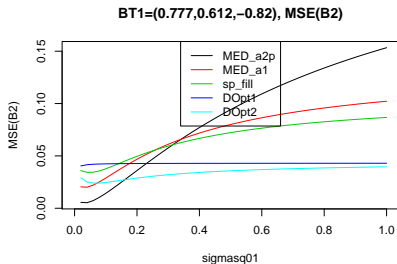
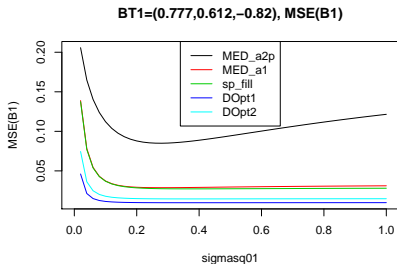
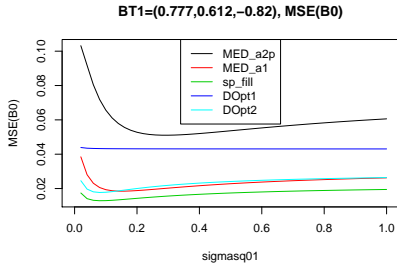
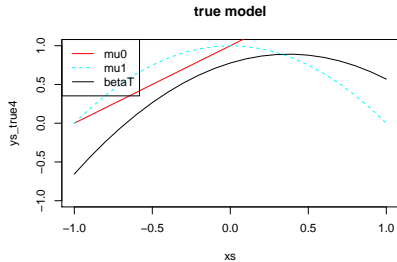
MSEs, BT(3)

betaTs[,3]

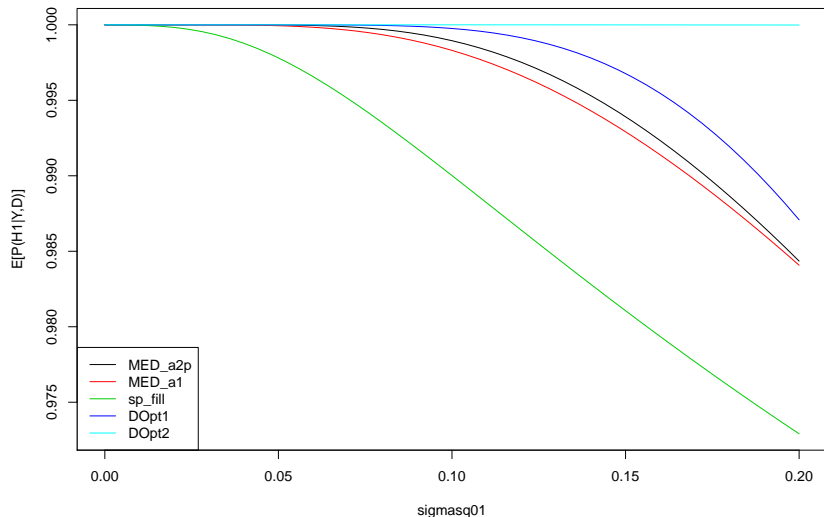


MSEs, BT(4)

betaTs[,4]

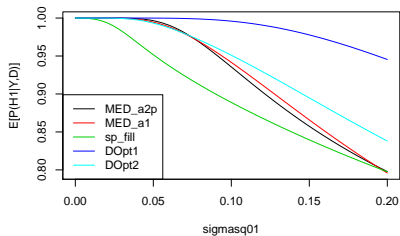


Example 4 $E[P(H1|Y,D, \text{beta}T=\mu1)]$

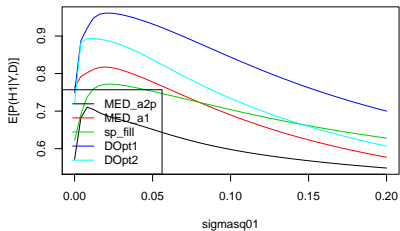


Example 4 $E[P(H1|Y,D)]$

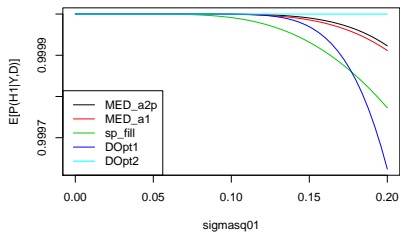
BT1=(0.72,-0.115,-0.221)



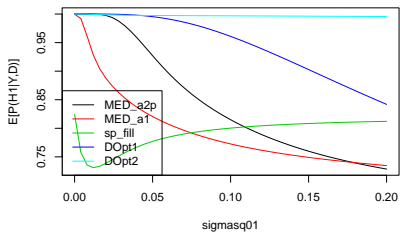
BT2=(1.035,0.065,-0.142)



BT3=(1.23,-0.633,-1.343)



BT4=(0.777,0.612,-0.82)



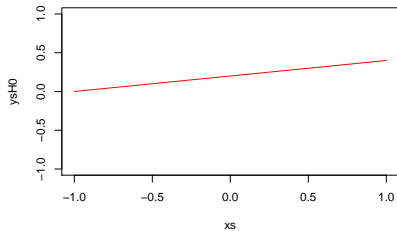
Example 4.1

Example 4.1 Parameters

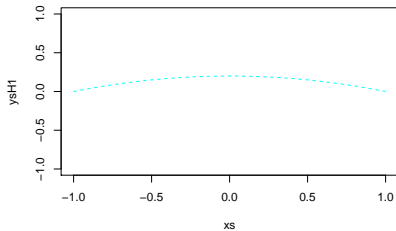
```
# Priors
sigmasq01 = 0.01
mu0 = c(0.2, 0.2)
V0 = diag(rep(sigmasq01,length(mu0)))
mu1 = c(0.2, 0, -0.2)
V1 = diag(rep(sigmasq01,length(mu1)))
sigmasq = 1
betaT = mu1
```

Example 4.1 Plots

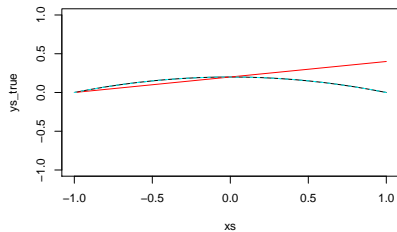
H0 model



H1 model

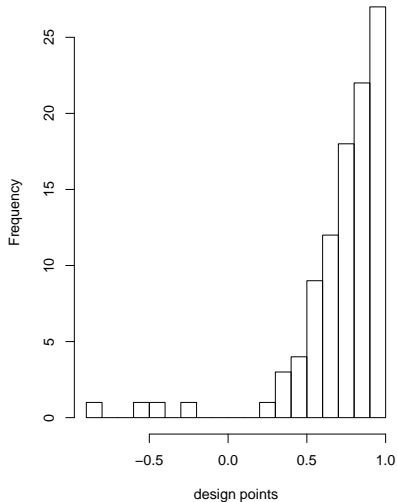


true model

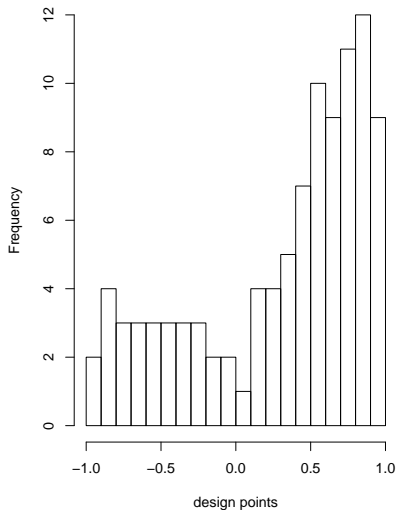


Example 4.1 MEDs

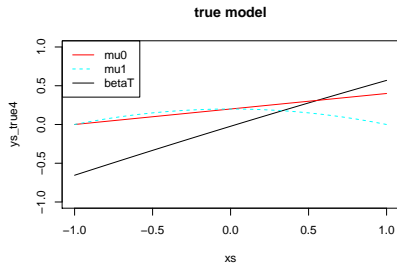
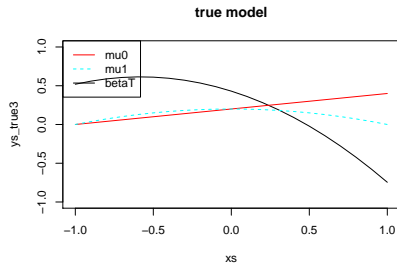
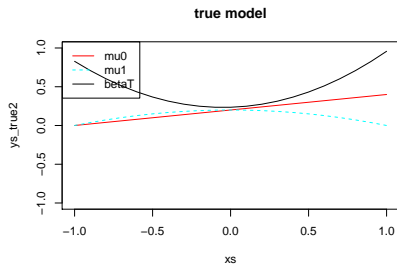
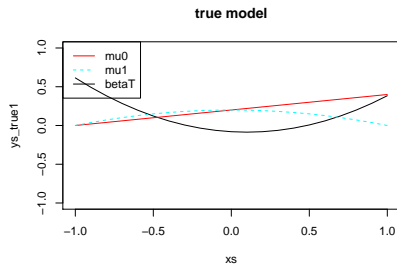
MED, $\alpha = 2p$



MED, $\alpha = 1$

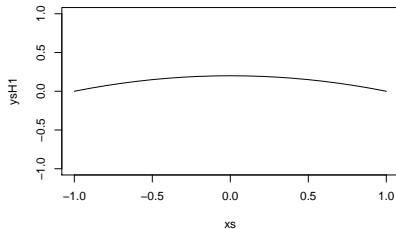


Example 4.1, Different BetaT

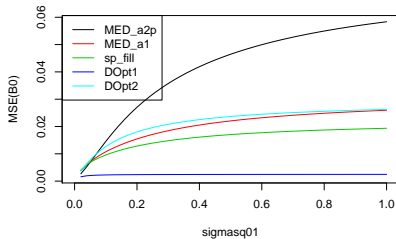


MSEs, BT = μ_1

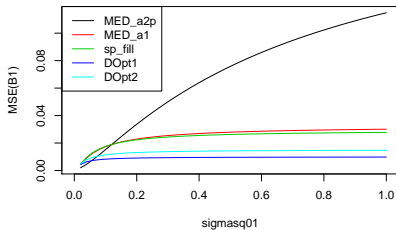
true model = μ_1



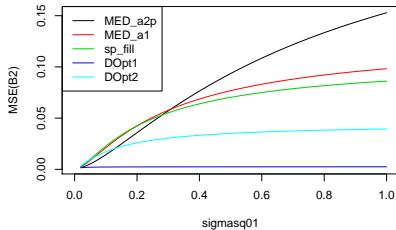
BT=(0.2,0,-0.2), MSE(B0)



BT=(0.2,0,-0.2), MSE(B1)



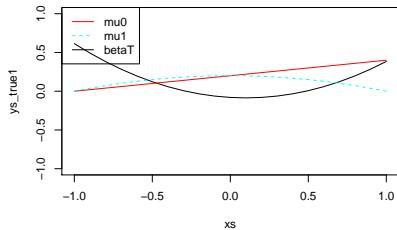
BT=(0.2,0,-0.2), MSE(B2)



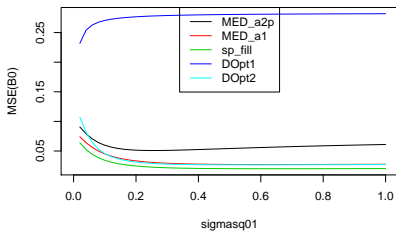
MSEs, BT(1)

betaTs[,1]

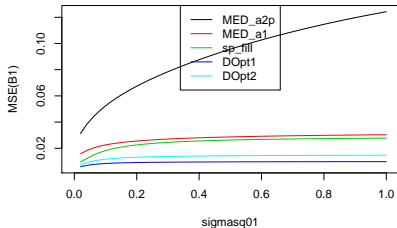
true model



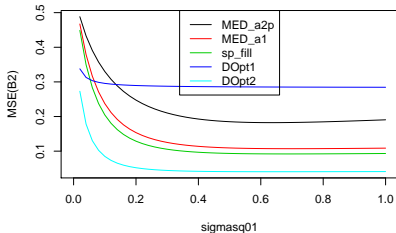
BT1=(-0.08,-0.115,0.579), MSE(B0)



BT1=(-0.08,-0.115,0.579), MSE(B1)



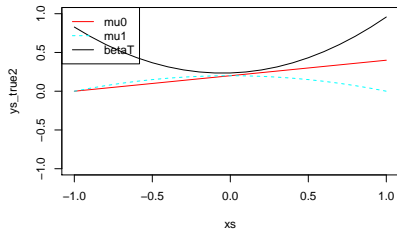
BT1=(-0.08,-0.115,0.579), MSE(B2)



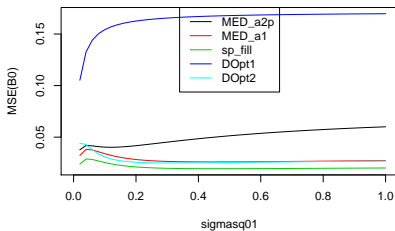
MSEs, BT(2)

betaTs[,2]

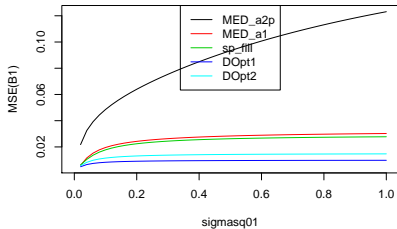
true model



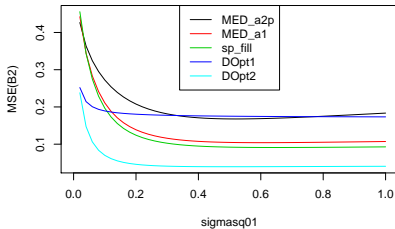
BT1=(0.235,0.065,0.658), MSE(B0)



BT1=(0.235,0.065,0.658), MSE(B1)

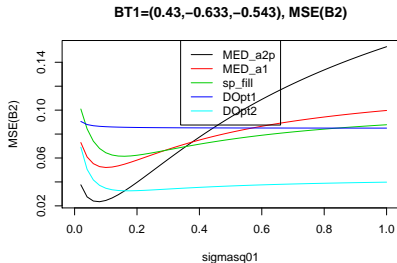
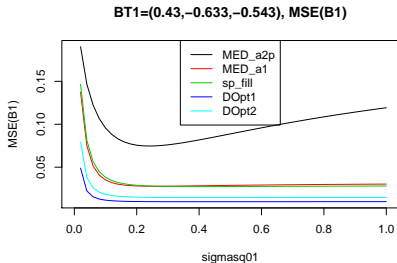
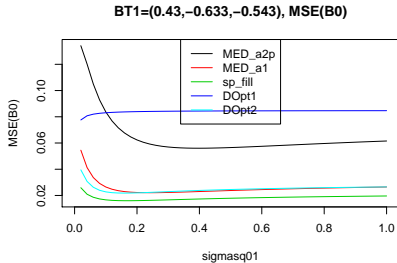
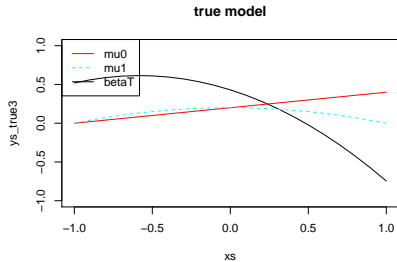


BT1=(0.235,0.065,0.658), MSE(B2)



MSEs, BT(3)

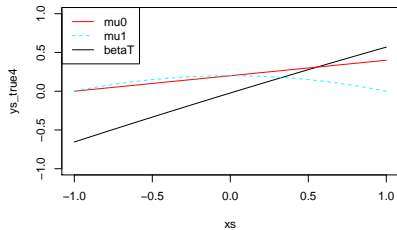
betaTs[,3]



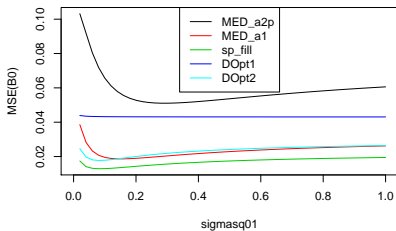
MSEs, BT(4)

betaTs[,4]

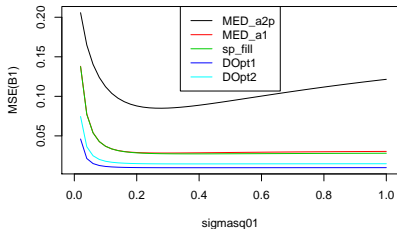
true model



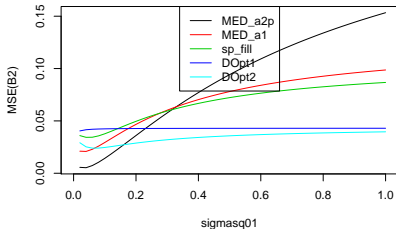
BT1=(-0.023,0.612,-0.02), MSE(B0)



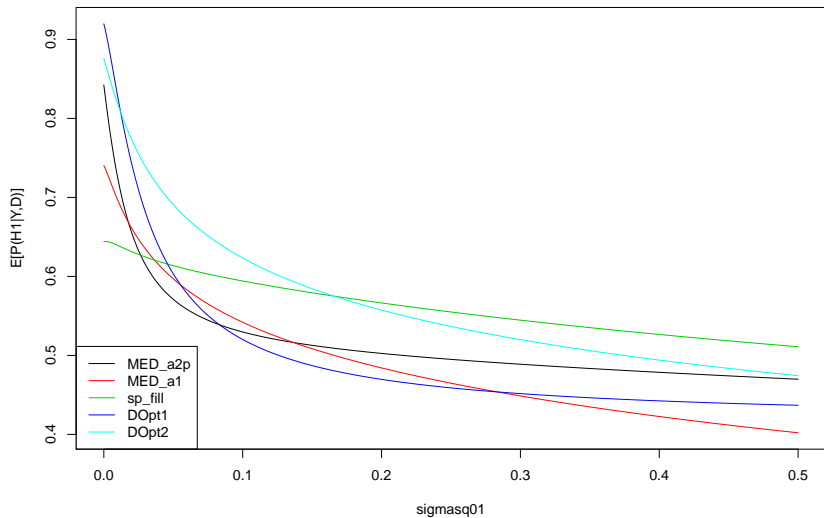
BT1=(-0.023,0.612,-0.02), MSE(B1)



BT1=(-0.023,0.612,-0.02), MSE(B2)

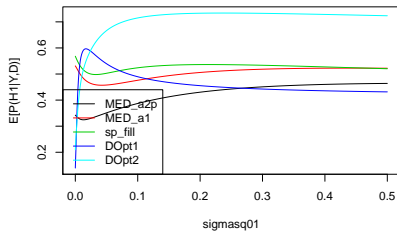


Example 4.1 $E[P(H1|Y,D, \text{betaT}=\mu1)]$

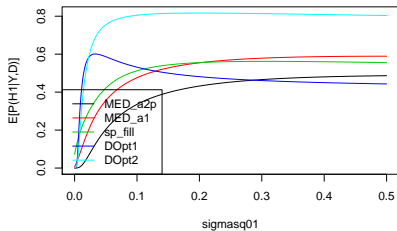


Example 4.1 $E[P(H1|Y,D)]$

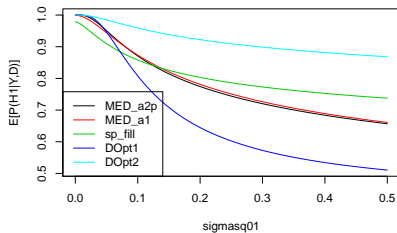
BT1=(-0.08,-0.115,0.579)



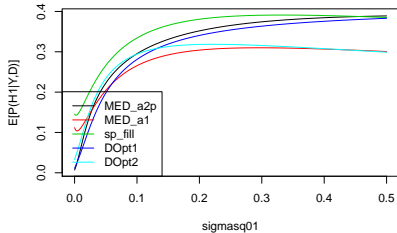
BT2=(0.235,0.065,0.658)



BT3=(0.43,-0.633,-0.543)



BT4=(-0.023,0.612,-0.02)



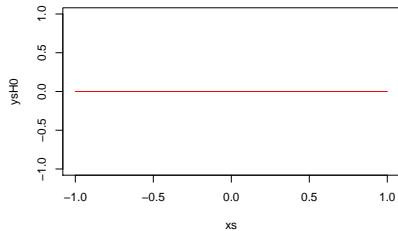
Example 5

Example 5 Parameters

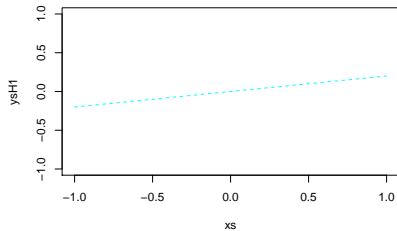
```
# Priors
sigmasq01 = 0.01
mu0 = c(0, 0)
V0 = diag(rep(sigmasq01,length(mu0)))
mu1 = c(0, 0.2)
V1 = diag(rep(sigmasq01,length(mu1)))
sigmasq = 1
betaT = mu1
```

Example 5 Plots

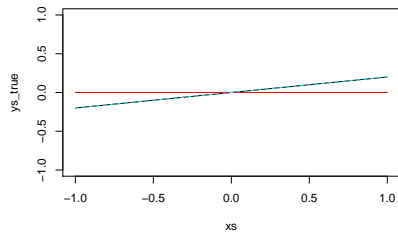
H0 model



H1 model

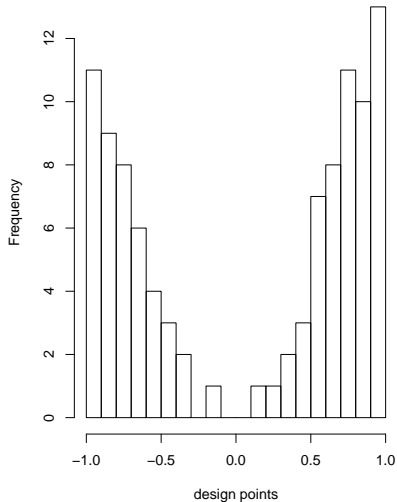


true model

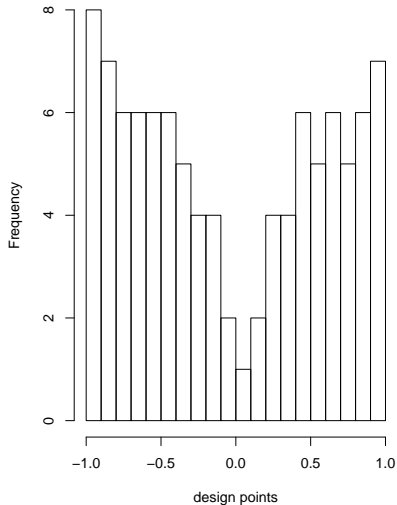


Example 5 MEDs

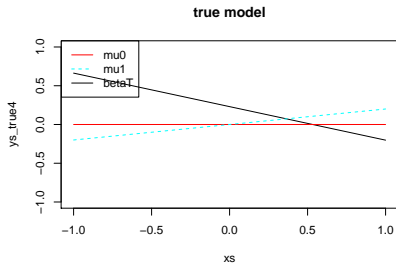
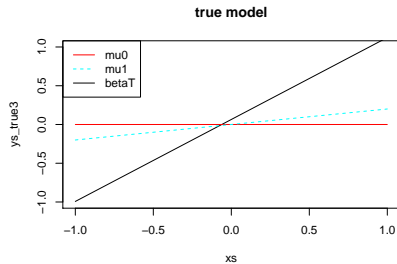
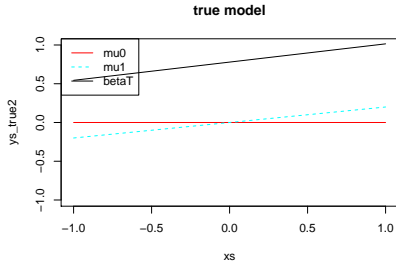
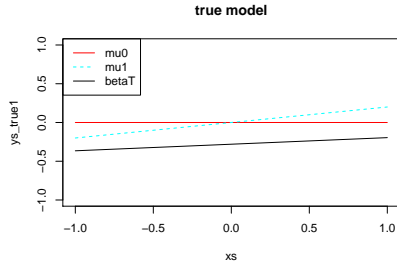
MED, $\alpha = 2p$



MED, $\alpha = 1$

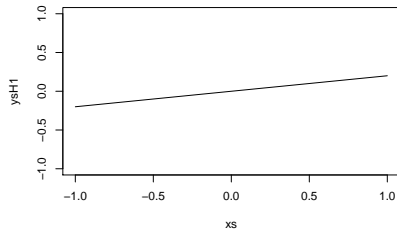


Example 5, Different BetaT

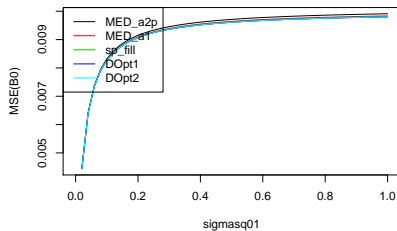


MSEs, $BT = \mu_1$

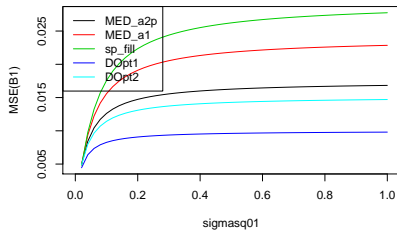
true model = μ_1



$BT=(0,0.2)$, $MSE(B_0)$

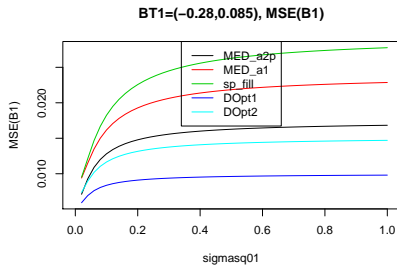
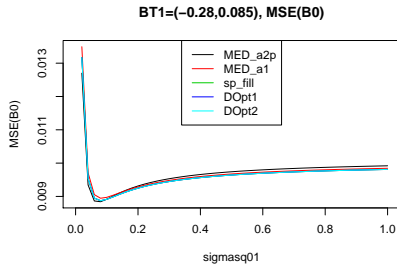
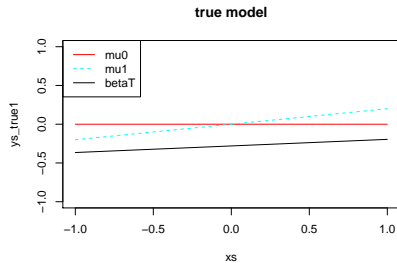


$BT=(0,0.2)$, $MSE(B_1)$



MSEs, BT(1)

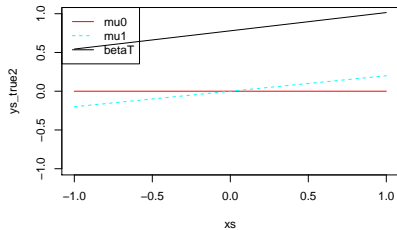
betaTs[,1]



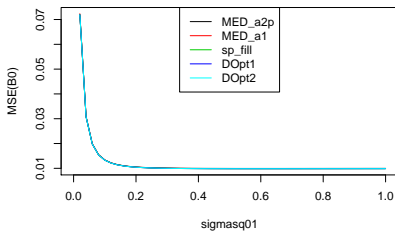
MSEs, BT(2)

betaTs[,2]

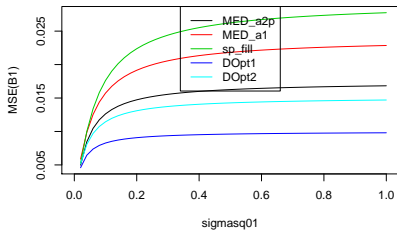
true model



BT1=(0.779,0.235), MSE(B0)

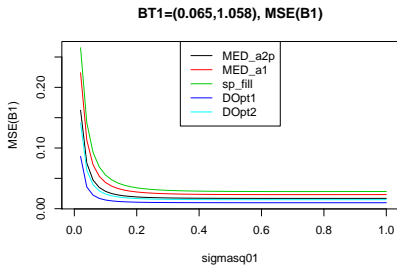
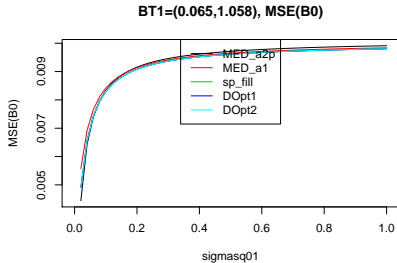
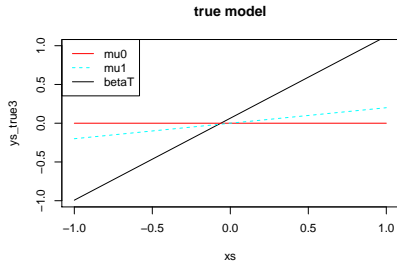


BT1=(0.779,0.235), MSE(B1)



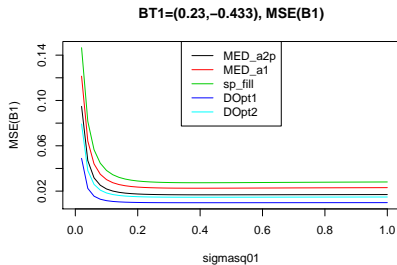
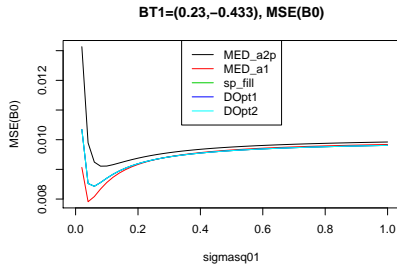
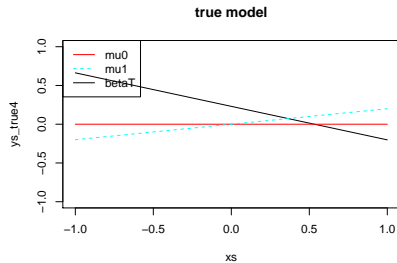
MSEs, BT(3)

betaTs[,3]

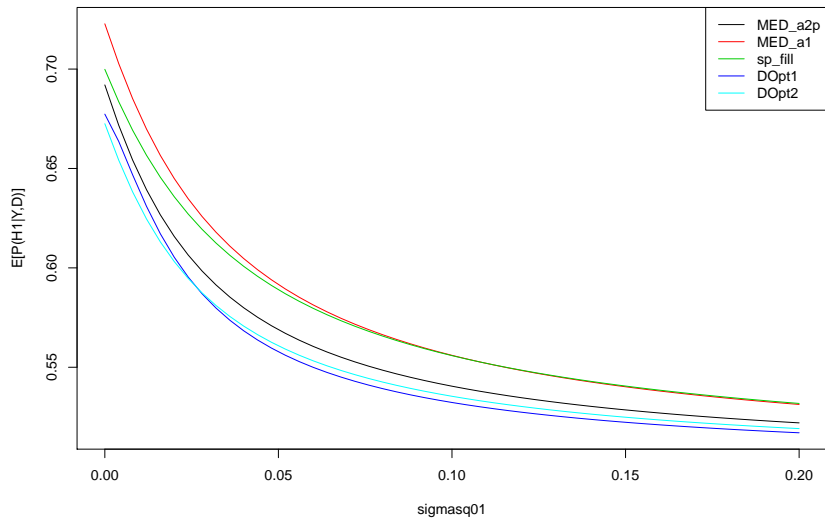


MSEs, BT(4)

betaTs[,4]

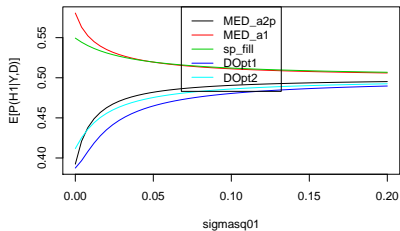


Example 5 $E[P(H1|Y,D, \text{beta}T=\mu1)]$

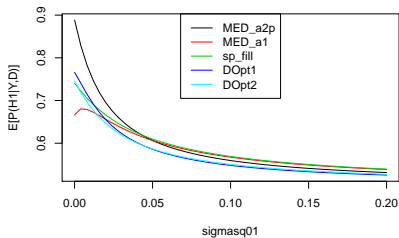


Example 5 $E[P(H1|Y,D)]$

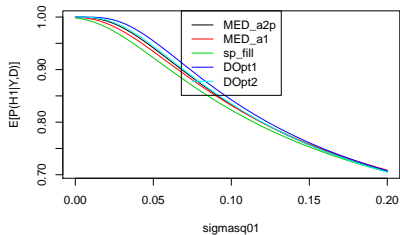
BT1=(-0.28,0.085)



BT2=(0.779,0.235)



BT3=(0.065,1.058)



BT4=(0.23,-0.433)

