MSE of Posterior Mean

Kristyn Pantoja

9/13/2019

Bias = 0

Bias != 0

The Main Question: Is it a problem that the squared differences as number of simulations, J, doesn't decrease monotonically for some parameters Beta_i?

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
$y \sim N(X\beta_T, \sigma^2 I)$

Empirical MSE converges to Closed Form MSE

- For each parameter β_i in $\beta = (\beta_1, \ldots, \beta_p)^T$, I calculated the closed form MSE and compare it to the mean empirical MSE.
- First, I consider the case where the bias should be 0, i.e.
  - for the null linear model $H_0 : E[\beta] = \mu_0$ where $\mu_0 \in \mathbb{R}^3$, I choose $\beta_T = \mu_0$
  - and likewise for the alternative quadratic model $H_1 : E[\beta] = \mu_1$ with $\mu_1 \in \mathbb{R}^5$, I choose $\beta_T = \mu_0$.
- I look at the values of the closed form MSE and the mean empirical MSE for each parameter β_i, estimated from J simulations. Call them $MSE(\beta_i)$ and $MSE_{MC,J}(\beta_i)$ respectively.
- I also see if, for simulations from 100 to 5000, that the squared difference between the closed from MSE and the mean empirical MSE for β_i goes to 0, i.e. $(MSE(\beta_i) - MSE_{MC,J}(\beta_i))^2 \to 0$ as J gets larger.
- And lastly, that the sum of squared differences betwen the closed from MSE and the mean empirical MSE for β_i over all β_i goes to 0, i.e. $\sum_{i=1}^{p}(MSE(\beta_i) - MSE_{MC,J}(\beta_i))^2 \to 0$
  - note that this is implied if $(MSE(\beta_i) - MSE_{MC,J}(\beta_i))^2 \to 0 \; \forall i$

Bias $= 0$

When Bias = 0

First consider the cases where the Bias^2 term should be 0, i.e. $\beta_T = \mu_\ell$, for $\ell = \{0, 1\}$.

Here, the number of simulations, J, is 5000.

```r
# evaluate : if they look similar
# H0 : mu0 = betaT = c(0, 0, 0)
space_closedMSE_v1$closedMSEH0$MSE_postmean
```

```
## [1] 0.0002279109 0.0003902575 0.0003902575
```

```r
space_EmpMSE_v1$expEmpMSEH0
```

```
## [1] 0.0002260150 0.0003752143 0.0003946635
```

```r
# H1 : mu1 = betaT = c(0, 0.2, -0.2, 0.2, -0.2)
space_closedMSE_v1$closedMSEH1$MSE_postmean
```

```
## [1] 0.0002084643 0.0006072928 0.0006021624 0.0006072928 0.0006021624
```

```r
space_EmpMSE_v1$expEmpMSEH1
```

```
## [1] 0.0002047821 0.0005730523 0.0005832644 0.0006037251 0.0006019561
```

Also note the bias terms are 0 in the closed form MSE for H_0 and H_1:

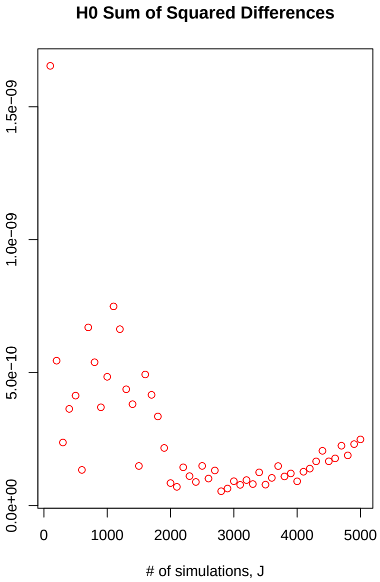
```r
# H0 : mu0 = betaT = c(0, 0, 0)
space_closedMSE_v1$closedMSEH0$biassq_term
```

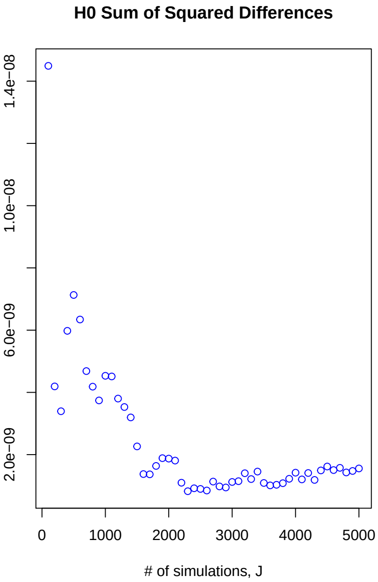
```
##      [,1]
## [1,]    0
## [2,]    0
## [3,]    0
```

```r
# H1 : mu1 = betaT = c(0, 0.2, -0.2, 0.2, -0.2)
space_closedMSE_v1$closedMSEH1$biassq_term
```

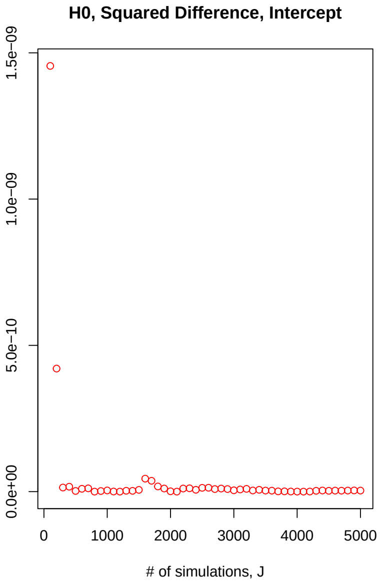
```
##              [,1]
## [1,] 3.900008e-33
## [2,] 0.000000e+00
## [3,] 0.000000e+00
## [4,] 0.000000e+00
## [5,] 6.938894e-18
```

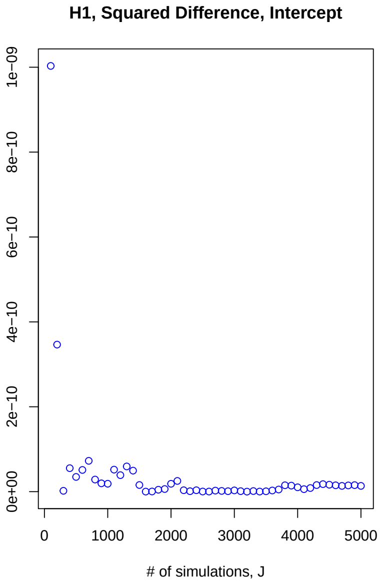
Sum of Squared Differences



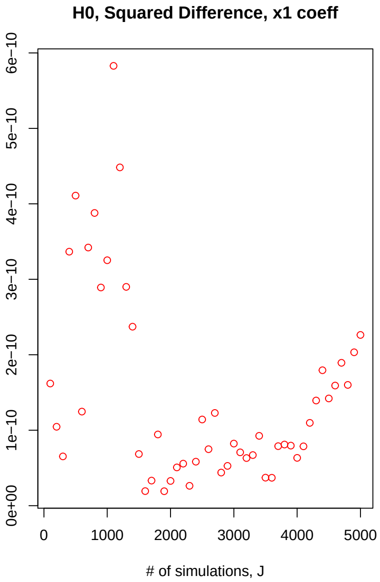


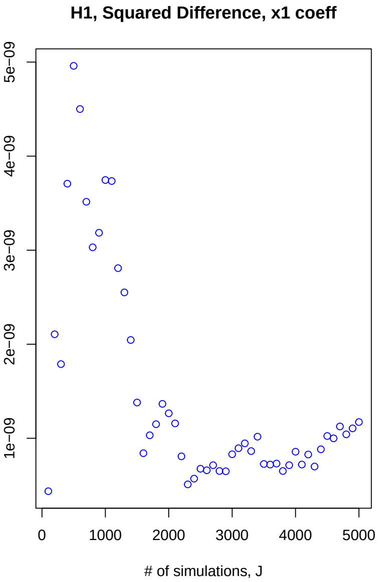
Squared Differences Beta0 (Intercept)



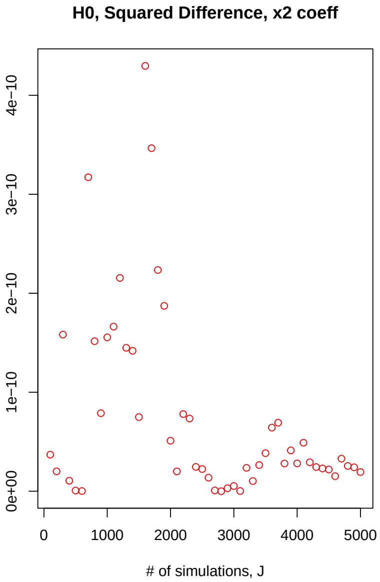


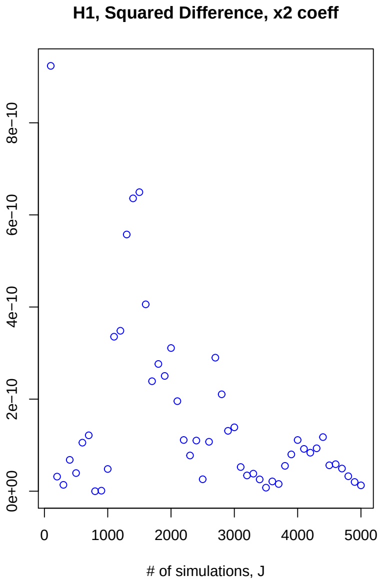
Squared Differences for Linear Term for x1



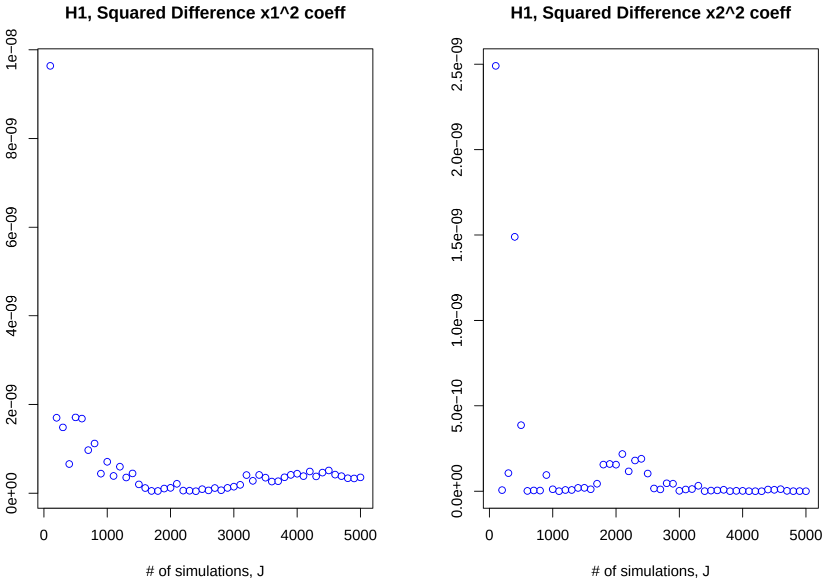


Squared Differences for Linear Term for x2





Squared Differences for Quadratic Terms x1, x2 in H1



Bias $!=$ 0

When Bias != 0

Here, for assuming a linear model as in H_0, I set $\beta_T = (0, 0.7, 0.7)$ and for assuming a quadratic model as in H_1, I set $\beta_T = (0, 0.7, -0.7, 0.7 - 0.7)$.

```
# evaluate : if they look similar
# H0 : mu0 = c(0, 0, 0) != betaT
space_closedMSE_v2$closedMSEH0$MSE_postmean
```

```
## [1] 0.003748783 0.003981899 0.003981899
```

```
space_EmpMSE_v2$expEmpMSEH0
```

```
## [1] 0.003734785 0.003930671 0.003996846
```

```
# H1 : mu1 = c(0, 0.2, -0.2, 0.2, -0.2) != betaT
space_closedMSE_v2$closedMSEH1$MSE_postmean
```

```
## [1] 0.01243412 0.12754542 0.12142891 0.12754542 0.12142891
```

```
space_EmpMSE_v2$expEmpMSEH1
```

```
## [1] 0.01239557 0.12729890 0.12141115 0.12723686 0.12103288
```

As expected, the bias terms are nonzero in the closed form MSE for H_0 and H_1 in these cases:

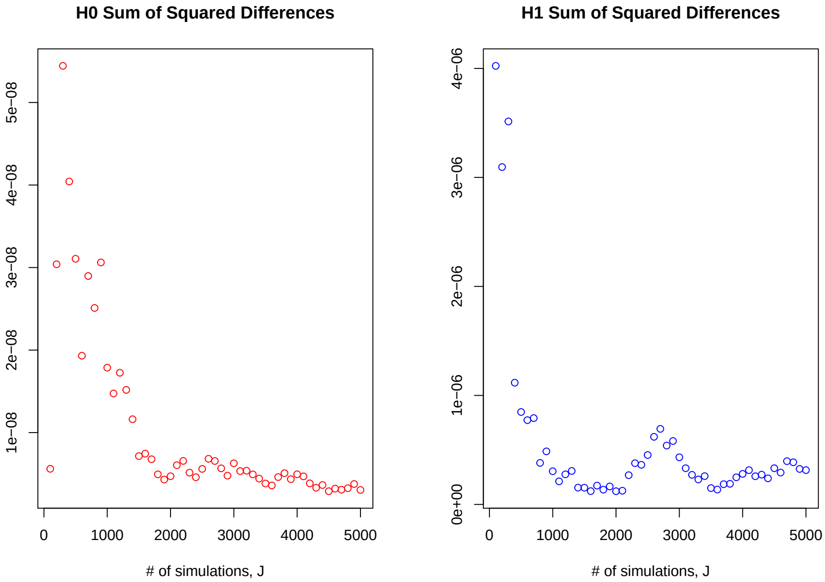
```
# H0 : mu0 = betaT = c(0, 0, 0)
space_closedMSE_v2$closedMSEH0$biassq_term
```

```
##             [,1]
## [1,] 0.003520872
## [2,] 0.003591641
## [3,] 0.003591641
```

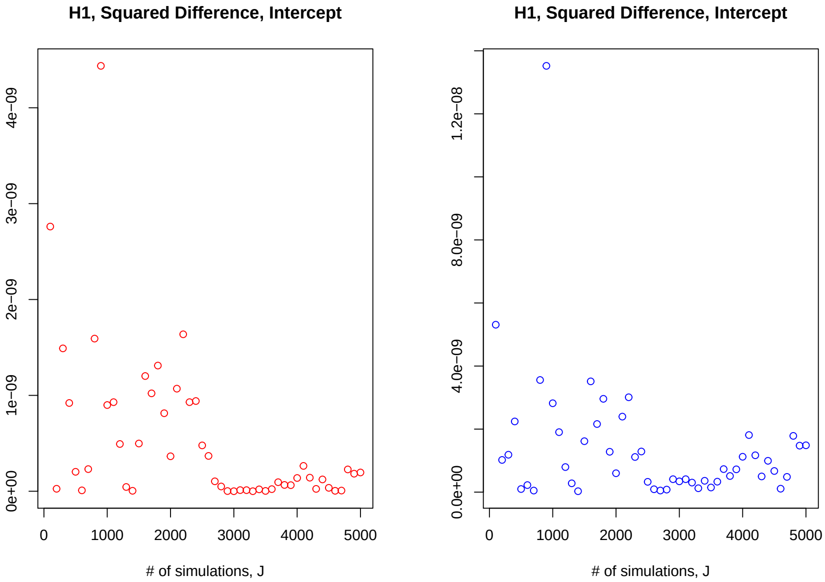
```
# H1 : mu1 = betaT = c(0, 0.2, -0.2, 0.2, -0.2)
space_closedMSE_v2$closedMSEH1$biassq_term
```

```
##            [,1]
## [1,] 0.01222565
## [2,] 0.12693813
## [3,] 0.12082675
## [4,] 0.12693813
## [5,] 0.12082675
```

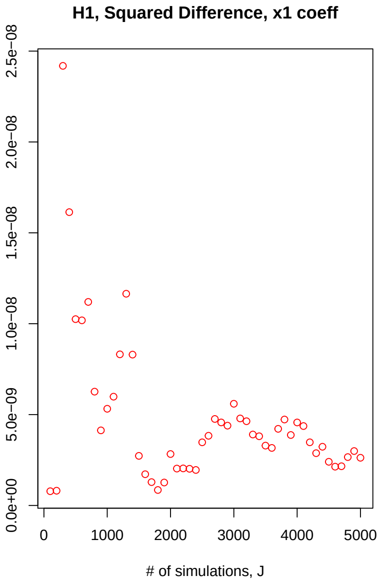
Sum of Squared Differences

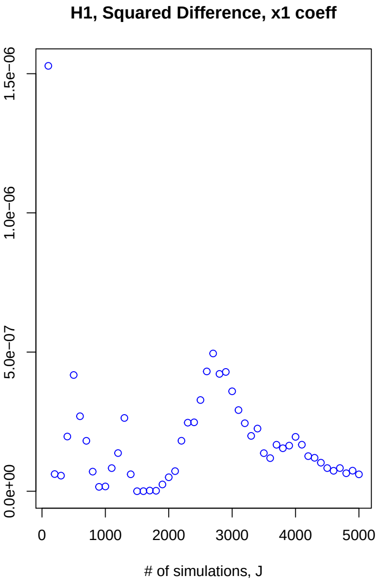


Squared Differences Beta0 (Intercept)

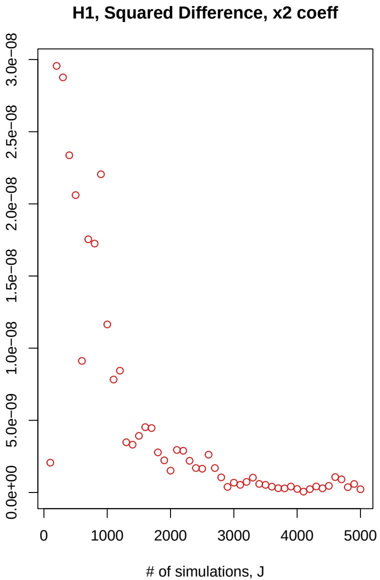


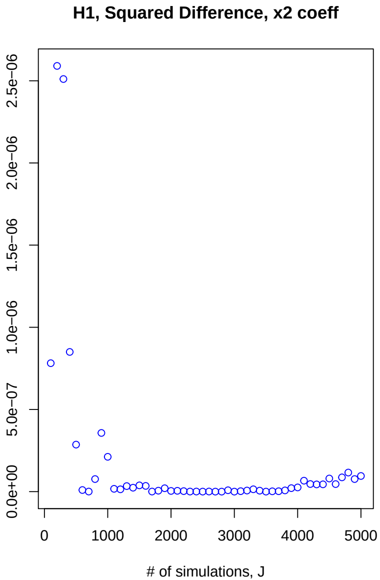
Squared Differences for Linear Term for x1



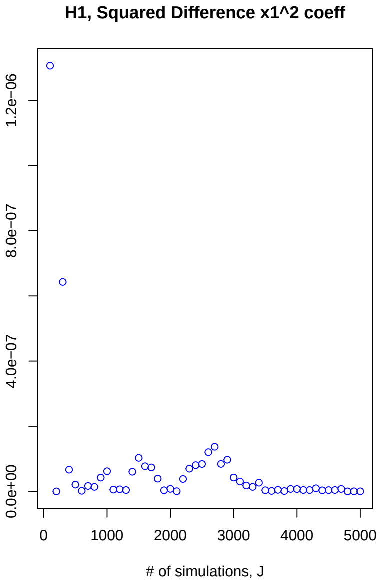


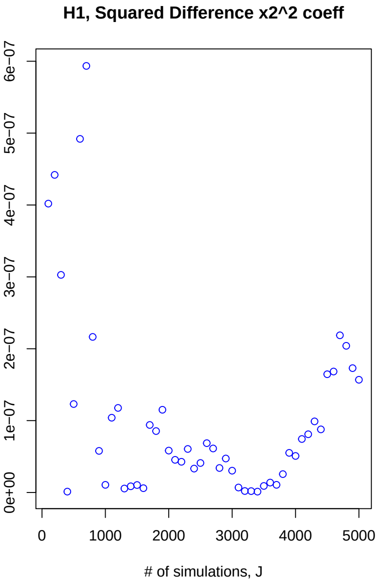
Squared Differences for Linear Term for x2





Squared Differences for Quadratic Terms x1, x2 in H1





The Main Question: Is it a problem that the squared differences as number of simulations, J, doesn't decrease monotonically for some parameters Beta_i?