Posterior Probabilities of Hypotheses

Kristyn Pantoja

9/10/2019

Closed Form MSE of Posterior Mean

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

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

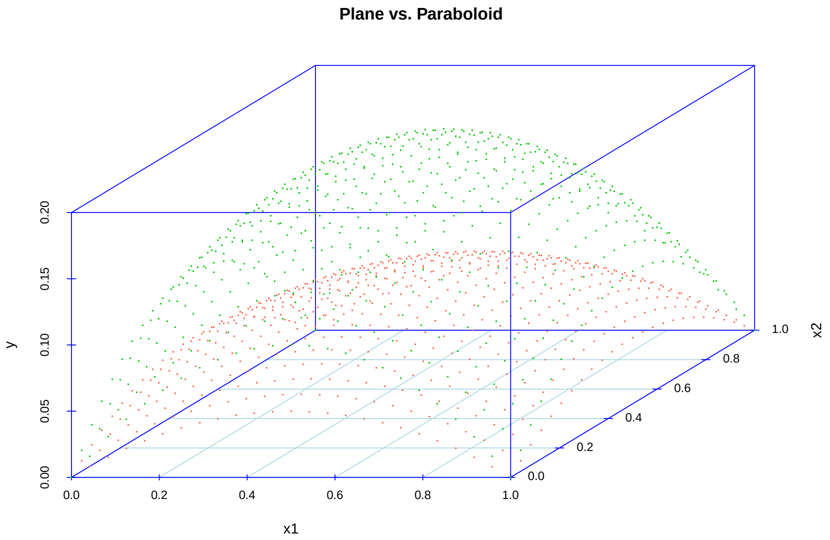
where

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

and $\Sigma_B = Var[\beta|Y] = \sigma^2(X^T X + \sigma^2 V^{-1})^{-1}$

MSE

- MSE to see the benefit of including space-filling in a design for parameter estimation: robust to model misspecification
  - We expect that the D-Optimal design will perform worse than the MEDs and space-filling designs.
- $H_0 : \beta \sim N(\mu_0, V_0)$
  - $y = \beta_0 + x_1\beta_1 + x_2\beta_2 + \varepsilon_i$
  - $\mu_0 = (0, 0, 0)^T$
- $H_1 : \beta \sim N(\mu_0, V_1)$
  - $y = \beta_0 + x_1\beta_1 + +x_1^2\beta_2 + x_2\beta_3 + +x_2^2\beta_4 + \varepsilon_i$
  - $\mu_1 = (0, -0.2, 0.2, -0.2, 0.2)^T$
- True Model: $\beta \sim N(\mu_T, V_1)$
  - $y = \beta_0 + x_1\beta_1 + +x_1^2\beta_2 + x_2\beta_3 + +x_2^2\beta_4 + \varepsilon_i$
  - $\mu_T = (0, -0.4, 0.4, -0.4, 0.4)^T$


Plane vs. Paraboloid
y
0.00
0.05
0.10
0.15
0.20
x1
0.0
0.2
0.4
0.6
0.8
1.0
x2
0.0
0.2
0.4
0.6
0.8
1.0

Results

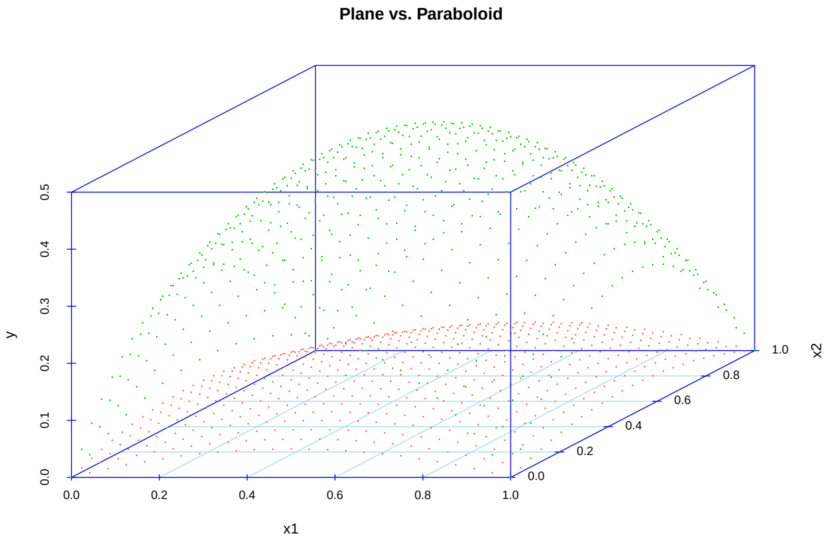
	MSE(B0)	MSE(B1)	MSE(B2)	MSE(B3)	MSE(B4)
k1	0.00449	0.04256	0.04261	0.04254	0.04254
k4	0.00447	0.04254	0.04256	0.04251	0.04256
k50	0.00446	0.04255	0.04254	0.04251	0.04262
fast	0.00452	0.04262	0.04270	0.04265	0.04273
dopt	0.00486	0.04245	0.04245	0.04245	0.04245
space	0.00471	0.04306	0.04317	0.04306	0.04317
random	0.00458	0.04275	0.04286	0.04276	0.04293
center	0.00306	0.04076	0.04019	0.04076	0.04019

MC Estimate of Empirical MSE

	E[MSE(B0)]	E[MSE(B1)]	E[MSE(B2)]	E[MSE(B3)]	E[MSE(B4)]
k1	0.00953	0.00338	0.00296	0.00343	0.00311
k4	0.00971	0.00350	0.00310	0.00338	0.00325
k50	0.00948	0.00344	0.00282	0.00332	0.00347
fast	0.00894	0.00363	0.00342	0.00408	0.00391
dopt	0.00479	0.00249	0.00249	0.00257	0.00257
space	0.00679	0.00653	0.00670	0.00691	0.00683
random	0.00908	0.00388	0.00489	0.00427	0.00508
center	0.00715	0.00179	0.00045	0.00179	0.00045

MSE, More Extreme Example

- MSE to the benefit of including space-filling in a design for parameter estimation: robust to model misspecification
  - We expect that the D-Optimal design will perform worse than the MEDs and space-filling designs.
- $H_0 : \beta \sim N(\mu_0, V_0)$
  - $y = \beta_0 + x_1\beta_1 + x_2\beta_2 + \varepsilon_i$
  - $\mu_0 = (0, 0, 0)^T$
- $H_1 : \beta \sim N(\mu_0, V_1)$
  - $y = \beta_0 + x_1\beta_1 + +x_1^2\beta_2 + x_2\beta_3 + +x_2^2\beta_4 + \varepsilon_i$
  - $\mu_1 = (0, -0.2, 0.2, -0.2, 0.2)^T$
- True Model: $\beta \sim N(\mu_T, V_1)$
  - $y = \beta_0 + x_1\beta_1 + +x_1^2\beta_2 + x_2\beta_3 + +x_2^2\beta_4 + \varepsilon_i$
  - $\mu_T = (0, -1, 1, -1, 1)^T$


Plane vs. Paraboloid
y
0.0
0.1
0.2
0.3
0.4
0.5
x1
0.0
0.2
0.4
0.6
0.8
1.0
x2
0.0
0.2
0.4
0.6
0.8
1.0

Results

	MSE(B0)	MSE(B1)	MSE(B2)	MSE(B3)	MSE(B4)
k1	0.00449	0.64256	0.64261	0.64254	0.64254
k4	0.00447	0.64254	0.64256	0.64251	0.64256
k50	0.00446	0.64255	0.64254	0.64251	0.64262
fast	0.00452	0.64262	0.64270	0.64265	0.64273
dopt	0.00486	0.64245	0.64245	0.64245	0.64245
space	0.00471	0.64306	0.64317	0.64306	0.64317
random	0.00458	0.64275	0.64286	0.64276	0.64293
center	0.00306	0.64076	0.64019	0.64076	0.64019

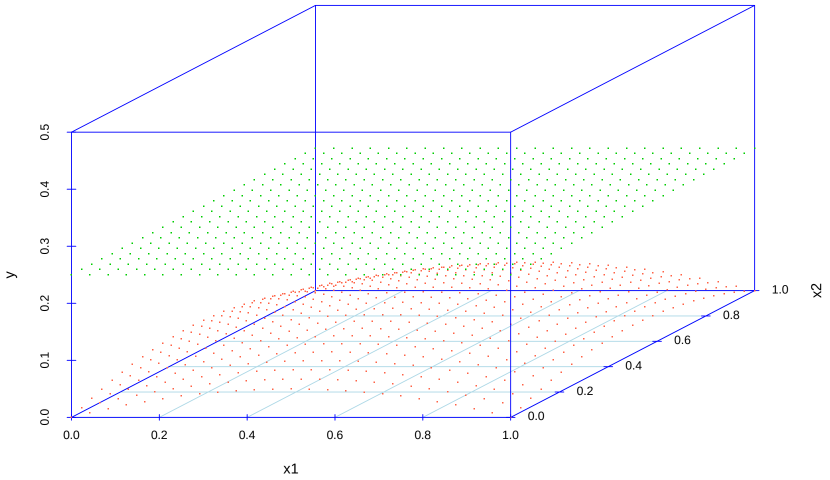
MC Estimate of Empirical MSE

	E[MSE(B0)]	E[MSE(B1)]	E[MSE(B2)]	E[MSE(B3)]	E[MSE(B4)]
k1	0.08274	0.01303	0.01071	0.01204	0.00911
k4	0.08414	0.01619	0.01273	0.01417	0.01349
k50	0.08418	0.01582	0.01033	0.01333	0.01499
fast	0.07566	0.01749	0.01707	0.02078	0.01958
dopt	0.00479	0.00249	0.00249	0.00257	0.00257
space	0.03649	0.05612	0.06223	0.05744	0.06245
random	0.07280	0.02047	0.03677	0.02670	0.03710
center	0.06387	0.01597	0.00399	0.01597	0.00399

A different Example: Plane

- Here, since model for D-Optimal design is not misspecified, it should have the best (lowest) MSE.
- $H_0 : \beta \sim N(\mu_0, V_0)$
  - $y = \beta_0 + x_1\beta_1 + x_2\beta_2 + \varepsilon_i$
  - $\mu_0 = (0, 0, 0)^T$
- $H_1 : \beta \sim N(\mu_0, V_1)$
  - $y = \beta_0 + x_1\beta_1 + +x_1^2\beta_2 + x_2\beta_3 + +x_2^2\beta_4 + \varepsilon_i$
  - $\mu_1 = (0, -0.2, 0.2, -0.2, 0.2)^T$
- True Model: $\beta \sim N(\mu_T, V_1)$
  - $y = \beta_0 + x_1\beta_1 + +x_1^2\beta_2 + x_2\beta_3 + +x_2^2\beta_4 + \varepsilon_i$
  - $\mu_T = (0.25, 0, 0, 0, 0)^T$

Plane vs. Paraboloid



Results

	MSE(B0)	MSE(B1)	MSE(B2)	MSE(B3)	MSE(B4)
k1	0.06699	0.04256	0.04261	0.04254	0.04254
k4	0.06697	0.04254	0.04256	0.04251	0.04256
k50	0.06696	0.04255	0.04254	0.04251	0.04262
fast	0.06702	0.04262	0.04270	0.04265	0.04273
dopt	0.06736	0.04245	0.04245	0.04245	0.04245
space	0.06721	0.04306	0.04317	0.04306	0.04317
random	0.06708	0.04275	0.04286	0.04276	0.04293
center	0.06556	0.04076	0.04019	0.04076	0.04019

MC Estimate of Empirical MSE

	E[MSE(B0)]	E[MSE(B1)]	E[MSE(B2)]	E[MSE(B3)]	E[MSE(B4)]
k1	0.08274	0.01303	0.01071	0.01204	0.00911
k4	0.08414	0.01619	0.01273	0.01417	0.01349
k50	0.08418	0.01582	0.01033	0.01333	0.01499
fast	0.07566	0.01749	0.01707	0.02078	0.01958
dopt	0.00479	0.00249	0.00249	0.00257	0.00257
space	0.03649	0.05612	0.06223	0.05744	0.06245
random	0.07280	0.02047	0.03677	0.02670	0.03710
center	0.06387	0.01597	0.00399	0.01597	0.00399