Summary

- Calculating the EVSI using non-parametric regression
  - Sampling future datasets
  - Summary Statistics
  - Using non-parametric regression
- Calculating the EVSI in `R`.
- Calculating the EVSI in `SAVI`

References

Strong et al. (2015) *Estimating the Expected Value of Sample Information Using the Probabilistic Sensitivity Analysis Sample: A Fast, Nonparametric Regression-Based Method.*

Expected Value of Sample Information (EVSI)

- We are considering collecting an additional dataset $\boldsymbol{X}$.

Expected Value of Sample Information (EVSI)

- We are considering collecting an additional dataset $\boldsymbol{X}$.
- What is *expected value* of collecting data $\boldsymbol{X}$?

Expected Value of Sample Information (EVSI)

- We are considering collecting an additional dataset $\boldsymbol{X}$.
- What is *expected value* of collecting data $\boldsymbol{X}$?

EVSI

$$\mathsf{EVSI} = \mathsf{E}_{\boldsymbol{X}}\left[\max_t \mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\{\mathsf{NB}(d,\boldsymbol{\theta})\}\right] - \max_t \mathsf{E}_{\boldsymbol{\theta}}\{\mathsf{NB}_t(\boldsymbol{\theta})\}$$

Expected Value of Sample Information (EVSI)

- As we did with the expression for EVPPI, we can rewrite EVSI as

$$\begin{aligned}\mathsf{EVSI} &= \mathsf{E}_{\boldsymbol{X}}\left[\max_t \mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\{\mathsf{NB}(d,\boldsymbol{\theta})\}\right] - \max_t \mathsf{E}_{\boldsymbol{\theta}}\{\mathsf{NB}_t(\boldsymbol{\theta})\} \\ &= \mathsf{E}_{\boldsymbol{X}}\left[\max_t \mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\{\mathsf{NB}(d,\boldsymbol{\theta})\}\right] - \max_t \mathsf{E}_{\boldsymbol{X}}\left[\mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\{\mathsf{NB}(d,\boldsymbol{\theta})\}\right]\end{aligned}$$

- This is helpful when it comes to Monte Carlo estimation

The PSA sample

- The probability sensitivity analysis (PSA) sample
  - Samples from $p(\boldsymbol{\theta})$ and corresponding net benefits

$$
\begin{array}{cccccc}
\theta_1^{(1)} & \cdots & \theta_p^{(1)} & \mathsf{NB}_1^{(1)} & \cdots & \mathsf{NB}_T^{(1)} \\
\theta_1^{(2)} & \cdots & \theta_p^{(2)} & \mathsf{NB}_1^{(2)} & \cdots & \mathsf{NB}_T^{(2)} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\theta_1^{(N)} & \cdots & \theta_p^{(N)} & \mathsf{NB}_1^{(N)} & \cdots & \mathsf{NB}_T^{(N)}
\end{array}
$$

- How do we sample from the marginal distribution $p(\boldsymbol{X})$?

Generating sample datasets

- How do we sample from the marginal distribution $p(\boldsymbol{X})$?
- Firstly, recognise that $p(\boldsymbol{X}, \boldsymbol{\theta}) = p(\boldsymbol{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})$

Generating sample datasets

- How do we sample from the marginal distribution $p(\boldsymbol{X})$?
- Firstly, recognise that $p(\boldsymbol{X}, \boldsymbol{\theta}) = p(\boldsymbol{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})$
- Then, recognise that if we sample from the joint distribution $p(\boldsymbol{\theta}, \boldsymbol{X})$ then the samples $\boldsymbol{x}$ drawn from this joint distribution have a marginal distribution $p(\boldsymbol{X})$

- How do we sample from the marginal distribution $p(\boldsymbol{X})$?
- Firstly, recognise that $p(\boldsymbol{X}, \boldsymbol{\theta}) = p(\boldsymbol{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})$
- Then, recognise that if we sample from the joint distribution $p(\boldsymbol{\theta}, \boldsymbol{X})$ then the samples $\boldsymbol{x}$ drawn from this joint distribution have a marginal distribution $p(\boldsymbol{X})$
- So we can do the following:

Generating sample datasets

- How do we sample from the marginal distribution $p(\boldsymbol{X})$?
- Firstly, recognise that $p(\boldsymbol{X}, \boldsymbol{\theta}) = p(\boldsymbol{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})$
- Then, recognise that if we sample from the joint distribution $p(\boldsymbol{\theta}, \boldsymbol{X})$ then the samples $\boldsymbol{x}$ drawn from this joint distribution have a marginal distribution $p(\boldsymbol{X})$
- So we can do the following:
  - Define study of interest to give us $p(\boldsymbol{X}|\boldsymbol{\theta})$

Generating sample datasets

- How do we sample from the marginal distribution $p(\boldsymbol{X})$?
- Firstly, recognise that $p(\boldsymbol{X}, \boldsymbol{\theta}) = p(\boldsymbol{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})$
- Then, recognise that if we sample from the joint distribution $p(\boldsymbol{\theta}, \boldsymbol{X})$ then the samples $\boldsymbol{x}$ drawn from this joint distribution have a marginal distribution $p(\boldsymbol{X})$
- So we can do the following:
  - Define study of interest to give us $p(\boldsymbol{X}|\boldsymbol{\theta})$
  - Generate samples $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(N)}$ from $p(\boldsymbol{\theta})$

Generating sample datasets

- How do we sample from the marginal distribution $p(\boldsymbol{X})$?
- Firstly, recognise that $p(\boldsymbol{X}, \boldsymbol{\theta}) = p(\boldsymbol{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})$
- Then, recognise that if we sample from the joint distribution $p(\boldsymbol{\theta}, \boldsymbol{X})$ then the samples $\boldsymbol{x}$ drawn from this joint distribution have a marginal distribution $p(\boldsymbol{X})$
- So we can do the following:
  - Define study of interest to give us $p(\boldsymbol{X}|\boldsymbol{\theta})$
  - Generate samples $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(N)}$ from $p(\boldsymbol{\theta})$
  - Generate a single sample of $\boldsymbol{x}$ from $p(\boldsymbol{x}|\boldsymbol{\theta}^{(i)})$ for each $i = 1, \ldots, N$

Generating sample datasets

- How do we sample from the marginal distribution $p(\boldsymbol{X})$?
- Firstly, recognise that $p(\boldsymbol{X}, \boldsymbol{\theta}) = p(\boldsymbol{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})$
- Then, recognise that if we sample from the joint distribution $p(\boldsymbol{\theta}, \boldsymbol{X})$ then the samples $\boldsymbol{x}$ drawn from this joint distribution have a marginal distribution $p(\boldsymbol{X})$
- So we can do the following:
  - Define study of interest to give us $p(\boldsymbol{X}|\boldsymbol{\theta})$
  - Generate samples $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(N)}$ from $p(\boldsymbol{\theta})$
  - Generate a single sample of $\boldsymbol{x}$ from $p(\boldsymbol{x}|\boldsymbol{\theta}^{(i)})$ for each $i = 1, \ldots, N$
  - This gives us a sample $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(N)}$ from $p(\boldsymbol{X})$

Generating sample datasets

- How do we sample from the marginal distribution $p(\boldsymbol{X})$?
- Firstly, recognise that $p(\boldsymbol{X}, \boldsymbol{\theta}) = p(\boldsymbol{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})$
- Then, recognise that if we sample from the joint distribution $p(\boldsymbol{\theta}, \boldsymbol{X})$ then the samples $\boldsymbol{x}$ drawn from this joint distribution have a marginal distribution $p(\boldsymbol{X})$
- So we can do the following:
  - Define study of interest to give us $p(\boldsymbol{X}|\boldsymbol{\theta})$
  - Generate samples $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(N)}$ from $p(\boldsymbol{\theta})$
  - Generate a single sample of $\boldsymbol{x}$ from $p(\boldsymbol{x}|\boldsymbol{\theta}^{(i)})$ for each $i = 1, \ldots, N$
  - This gives us a sample $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(N)}$ from $p(\boldsymbol{X})$
- Note that any single sample $\boldsymbol{x}$ is a sample dataset from the proposed study.

- How do we sample from the marginal distribution $p(\boldsymbol{X})$?
- Firstly, recognise that $p(\boldsymbol{X}, \boldsymbol{\theta}) = p(\boldsymbol{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})$
- Then, recognise that if we sample from the joint distribution $p(\boldsymbol{\theta}, \boldsymbol{X})$ then the samples $\boldsymbol{x}$ drawn from this joint distribution have a marginal distribution $p(\boldsymbol{X})$
- So we can do the following:
    - Define study of interest to give us $p(\boldsymbol{X}|\boldsymbol{\theta})$
    - Generate samples $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(N)}$ from $p(\boldsymbol{\theta})$
    - Generate a single sample of $\boldsymbol{x}$ from $p(\boldsymbol{x}|\boldsymbol{\theta}^{(i)})$ for each $i = 1, \ldots, N$
    - This gives us a sample $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(N)}$ from $p(\boldsymbol{X})$
- Note that any single sample $\boldsymbol{x}$ is a sample dataset from the proposed study.
- Each sample $\boldsymbol{x}$ may be a vector of values

- Define study of interest
- This gives us $p(\boldsymbol{X}|\boldsymbol{\theta})$
- For each row in the PSA, sample a dataset $\boldsymbol{x}^{(i)}$ from $p(\boldsymbol{X}|\boldsymbol{\theta}^{(i)})$

$$\begin{array}{cccccc}
\theta_1^{(1)} & \cdots & \theta_p^{(1)} & \mathsf{NB}_1^{(1)} & \cdots & \mathsf{NB}_T^{(1)} \\
\theta_1^{(2)} & \cdots & \theta_p^{(2)} & \mathsf{NB}_1^{(2)} & \cdots & \mathsf{NB}_T^{(2)} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\theta_1^{(N)} & \cdots & \theta_p^{(N)} & \mathsf{NB}_1^{(N)} & \cdots & \mathsf{NB}_T^{(N)}
\end{array}$$

- Define study of interest
- This gives us $p(\boldsymbol{X}|\boldsymbol{\theta})$
- For each row in the PSA, sample a dataset $\boldsymbol{x}^{(i)}$ from $p(\boldsymbol{X}|\boldsymbol{\theta}^{(i)})$

$$\begin{array}{ccccccc}
\theta_1^{(1)} & \cdots & \theta_p^{(1)} & \mathsf{NB}_1^{(1)} & \cdots & \mathsf{NB}_T^{(1)} & \boldsymbol{x}^{(1)} \\
\theta_1^{(2)} & \cdots & \theta_p^{(2)} & \mathsf{NB}_1^{(2)} & \cdots & \mathsf{NB}_T^{(2)} & \boldsymbol{x}^{(2)} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\theta_1^{(N)} & \cdots & \theta_p^{(N)} & \mathsf{NB}_1^{(N)} & \cdots & \mathsf{NB}_T^{(N)} & \boldsymbol{x}^{(N)}
\end{array}$$

Expected Value of Sample Information (EVSI)

EVSI

$$\mathsf{EVSI} = \mathsf{E}_{\boldsymbol{X}} \left[\max_t \mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}} \{ \mathsf{NB}(d, \boldsymbol{\theta}) \} \right] - \max_t \mathsf{E}_{\boldsymbol{X}} \left[\mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}} \{ \mathsf{NB}(d, \boldsymbol{\theta}) \} \right]$$

EVSI

$$\text{EVSI} = \mathsf{E}_{\boldsymbol{X}}\left[\max_t \mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\{\text{NB}(d,\boldsymbol{\theta})\}\right] - \max_t \mathsf{E}_{\boldsymbol{X}}\left[\mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\{\text{NB}(d,\boldsymbol{\theta})\}\right]$$

$\mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\{\text{NB}(d,\boldsymbol{\theta})\}$ potentially difficult

Treat as regression problem

EVSI

$$\mathsf{EVSI} = \mathsf{E}_{\boldsymbol{X}}\left[\max_t \mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\{\mathsf{NB}(d,\boldsymbol{\theta})\}\right] - \max_t \mathsf{E}_{\boldsymbol{X}}\left[\mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\{\mathsf{NB}(d,\boldsymbol{\theta})\}\right]$$

Treat as regression problem

EVSI

$$\mathsf{EVSI} = \mathsf{E}_{\boldsymbol{X}} \left[\max_t \mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}} \{ \mathsf{NB}(d, \boldsymbol{\theta}) \} \right] - \max_t \mathsf{E}_{\boldsymbol{X}} \left[\mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}} \{ \mathsf{NB}(d, \boldsymbol{\theta}) \} \right]$$

- Write the following

Regression equation

$$\begin{aligned} \mathsf{NB}_t(\boldsymbol{\theta}) &= \mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}} \{ \mathsf{NB}_t(\boldsymbol{\theta}) \} + \varepsilon \\ &= g_t(\boldsymbol{X}) + \varepsilon \end{aligned}$$

Treat as regression problem

EVSI

$$\text{EVSI} = \text{E}_{\boldsymbol{X}}\left[\max_t \text{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\{\text{NB}(d,\boldsymbol{\theta})\}\right] - \max_t \text{E}_{\boldsymbol{X}}\left[\text{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\{\text{NB}(d,\boldsymbol{\theta})\}\right]$$

- Write the following

Regression equation

$$\begin{aligned}\text{NB}_t(\boldsymbol{\theta}) &= \text{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\{\text{NB}_t(\boldsymbol{\theta})\} + \varepsilon \\ &= g_t(\boldsymbol{X}) + \varepsilon\end{aligned}$$

- Where $g_t(\boldsymbol{X})$ is some unknown *smooth* function of $\boldsymbol{X}$

Treat as regression problem

EVSI

$$\mathsf{EVSI} = \mathsf{E}_{\boldsymbol{X}}\left[\max_t \mathsf{E}_{\boldsymbol{\theta}\mid\boldsymbol{X}}\{\mathsf{NB}(d,\boldsymbol{\theta})\}\right] - \max_t \mathsf{E}_{\boldsymbol{X}}\left[\mathsf{E}_{\boldsymbol{\theta}\mid\boldsymbol{X}}\{\mathsf{NB}(d,\boldsymbol{\theta})\}\right]$$

- Write the following

Regression equation

$$\begin{aligned}\mathsf{NB}_t(\boldsymbol{\theta}) &= \mathsf{E}_{\boldsymbol{\theta}\mid\boldsymbol{X}}\{\mathsf{NB}_t(\boldsymbol{\theta})\} + \varepsilon \\ &= g_t(\boldsymbol{X}) + \varepsilon\end{aligned}$$

- Where $g_t(\boldsymbol{X})$ is some unknown *smooth* function of $\boldsymbol{X}$
- Treat as non-parametric regression of $\mathsf{NB}_t(\boldsymbol{\theta})$ on $\boldsymbol{X}$

Treat as regression problem

EVSI

$$\text{EVSI} = \mathsf{E}_{\boldsymbol{X}}\left[\max_t \mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\{\text{NB}(d,\boldsymbol{\theta})\}\right] - \max_t \mathsf{E}_{\boldsymbol{X}}\left[\mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\{\text{NB}(d,\boldsymbol{\theta})\}\right]$$

- Write the following

Regression equation

$$\begin{aligned}\text{NB}_t(\boldsymbol{\theta}) &= \mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\{\text{NB}_t(\boldsymbol{\theta})\} + \varepsilon \\ &= g_t(\boldsymbol{X}) + \varepsilon\end{aligned}$$

- Where $g_t(\boldsymbol{X})$ is some unknown *smooth* function of $\boldsymbol{X}$
- Treat as non-parametric regression of $\text{NB}_t(\boldsymbol{\theta})$ on $\boldsymbol{X}$
- If $\boldsymbol{X}$ is vector, summarise $S(\boldsymbol{X})$ and regress on $S(\boldsymbol{X})$

PSA plus datasets

$$
\begin{array}{ccccccc}
\theta_1^{(1)} & \cdots & \theta_p^{(1)} & \mathsf{NB}_1^{(1)} & \cdots & \mathsf{NB}_T^{(1)} & \boldsymbol{x}^{(1)} \\
\theta_1^{(2)} & \cdots & \theta_p^{(2)} & \mathsf{NB}_1^{(2)} & \cdots & \mathsf{NB}_T^{(2)} & \boldsymbol{x}^{(2)} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\theta_1^{(N)} & \cdots & \theta_p^{(N)} & \mathsf{NB}_1^{(N)} & \cdots & \mathsf{NB}_T^{(N)} & \boldsymbol{x}^{(N)}
\end{array}
$$

PSA plus datasets

$$
\begin{array}{cccccccc}
\theta_1^{(1)} & \cdots & \theta_p^{(1)} & \mathsf{NB}_1^{(1)} & \cdots & \mathsf{NB}_T^{(1)} & \boldsymbol{x}^{(1)} & S(\boldsymbol{x}^{(1)}) \\
\theta_1^{(2)} & \cdots & \theta_p^{(2)} & \mathsf{NB}_1^{(2)} & \cdots & \mathsf{NB}_T^{(2)} & \boldsymbol{x}^{(2)} & S(\boldsymbol{x}^{(2)}) \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\theta_1^{(N)} & \cdots & \theta_p^{(N)} & \mathsf{NB}_1^{(N)} & \cdots & \mathsf{NB}_T^{(N)} & \boldsymbol{x}^{(N)} & S(\boldsymbol{x}^{(N)})
\end{array}
$$

Treat as regression problem

- For each dataset calculate scalar summary statistic $S(\boldsymbol{x}^{(i)})$
- Regress net benefits $\mathrm{NB}_t(\boldsymbol{\theta}^{(i)})$ on $S(\boldsymbol{x}^{(i)})$

[1] Strong M, Oakley JE, Brennan A, Breeze P. Estimating the Expected Value of Sample Information using the Probabilistic Sensitivity Analysis Sample. A Fast Non-Parametric Regression Based Method. *Medical Decision Making*. 2015;**35(5)**:570-583

Treat as regression problem

- For each dataset calculate scalar summary statistic $S(\boldsymbol{x}^{(i)})$
- Regress net benefits $\mathsf{NB}_t(\boldsymbol{\theta}^{(i)})$ on $S(\boldsymbol{x}^{(i)})$
- Repeat for each $t \in \{1, \ldots, T\}$

- Extract fitted values $\hat{g}_t\{S(\boldsymbol{x}^{(i)})\}$
- These are estimates of $\mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}^{(i)}}\{\mathsf{NB}_t(\boldsymbol{\theta})\}$

[1] Strong M, Oakley JE, Brennan A, Breeze P. Estimating the Expected Value of Sample Information using the Probabilistic Sensitivity Analysis Sample. A Fast Non-Parametric Regression Based Method. *Medical Decision Making*. 2015;**35(5)**:570-583

Treat as regression problem

- For each dataset calculate scalar summary statistic $S(\boldsymbol{x}^{(i)})$
- Regress net benefits $\mathsf{NB}_t(\boldsymbol{\theta}^{(i)})$ on $S(\boldsymbol{x}^{(i)})$
- Repeat for each $t \in \{1, \ldots, T\}$

- Extract fitted values $\hat{g}_t\{S(\boldsymbol{x}^{(i)})\}$
- These are estimates of $\mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}^{(i)}}\{\mathsf{NB}_t(\boldsymbol{\theta})\}$

- EVSI estimator is now just[1]

$$\frac{1}{N}\sum_{i=1}^{N}\max_t \hat{g}_t\{S(\boldsymbol{x}^{(i)})\} - \max_t \frac{1}{N}\sum_{i=1}^{N}\hat{g}_t\{S(\boldsymbol{x}^{(i)})\}$$

[1] Strong M, Oakley JE, Brennan A, Breeze P. Estimating the Expected Value of Sample Information using the Probabilistic Sensitivity Analysis Sample. A Fast Non-Parametric Regression Based Method. *Medical Decision Making*. 2015;**35(5)**:570-583

Example R code

```
library(mgcv)
model <- gam(INB ~ s(Sx))
g.hat <- fitted(model)
evsi <- mean(pmax(0, g.hat)) - max(0, mean(g.hat))
```

Empirical case study results

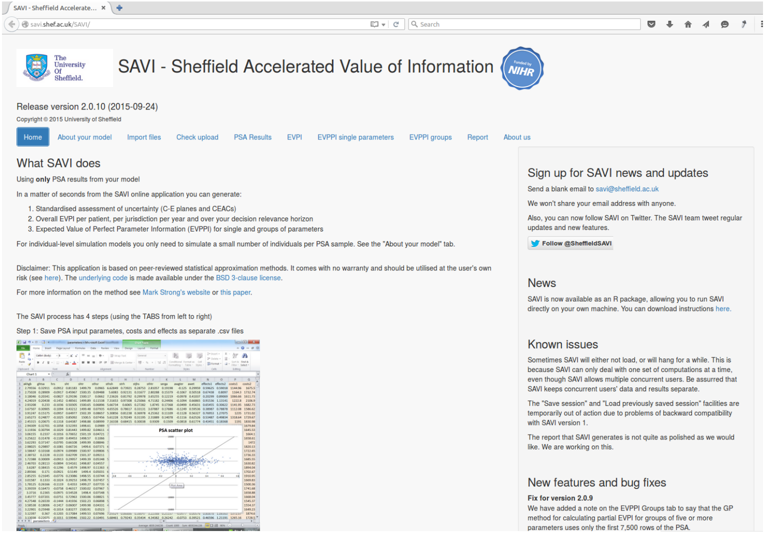
Sample sizes			Coefficient of variation	Computation time (s)
Outer	Inner	Total		
Two-level Monte Carlo method				
10^4	10^4	10^8	1.9%	4,456
10^5	10^4	10^9	0.6%	43,303
10^6	10^4	10^{10}	0.2%	424,686
Non-parametric regression method				
10^4	-	10^4	2.4%	0.1
10^5	-	10^5	0.8%	0.7
10^6	-	10^6	0.2%	8

Empirical case study results

Sample sizes			Coefficient of variation	Computation time (s)
Outer	Inner	Total		
Two-level Monte Carlo method				
10^4	10^4	10^8	1.9%	4,456
10^5	10^4	10^9	0.6%	43,303
10^6	10^4	10^{10}	0.2%	424,686
Non-parametric regression method				
10^4	-	10^4	2.4%	0.1
10^5	-	10^5	0.8%	0.7
10^6	-	10^6	0.2%	8

50,000 times speed up!

Example in SAVI



Summary

- Fast and efficient method for EVSI
- Only the PSA sample required
- No need to run the model again
- ‘Black box’ Value of Information