Fast approximations for the Expected Value of Partial Perfect Information using R-INLA

Anna Heath[1]

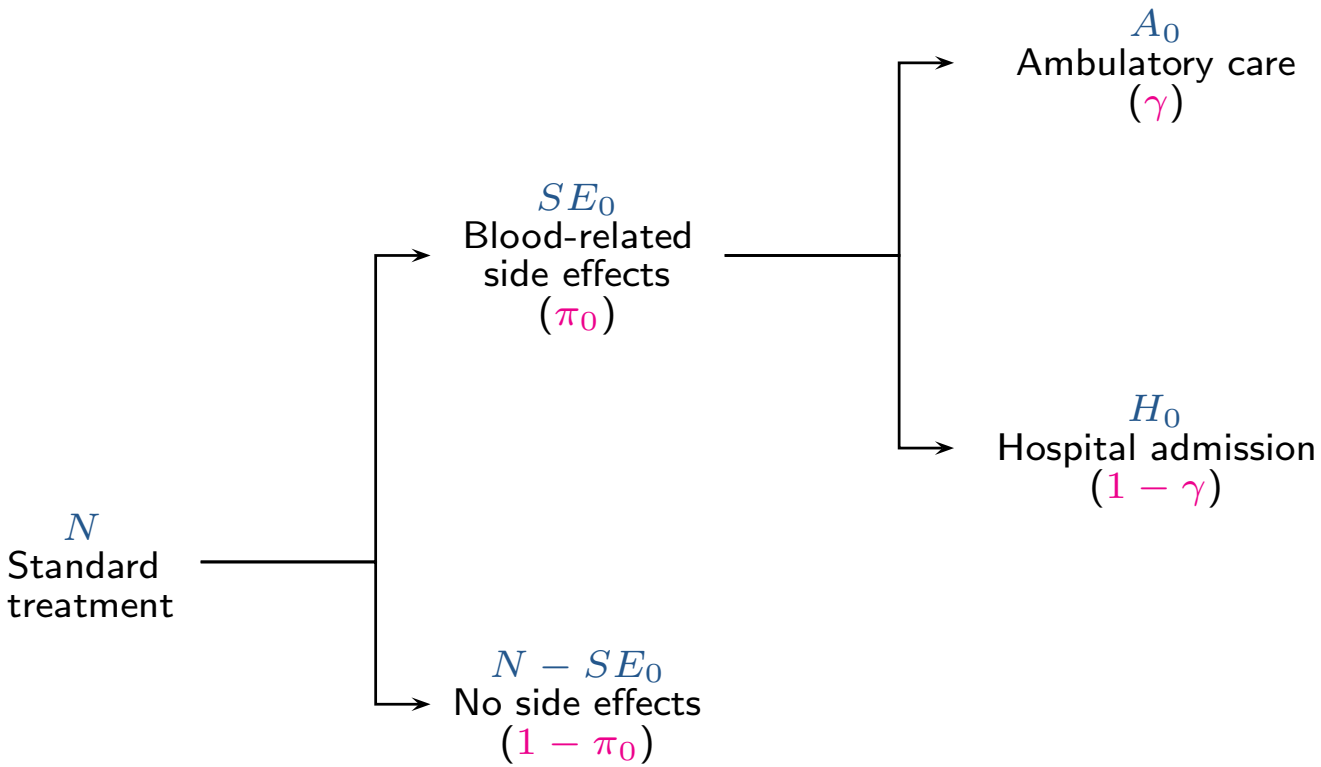
[1] Department of Statistical Science, University College London

22 May 2015

Outline

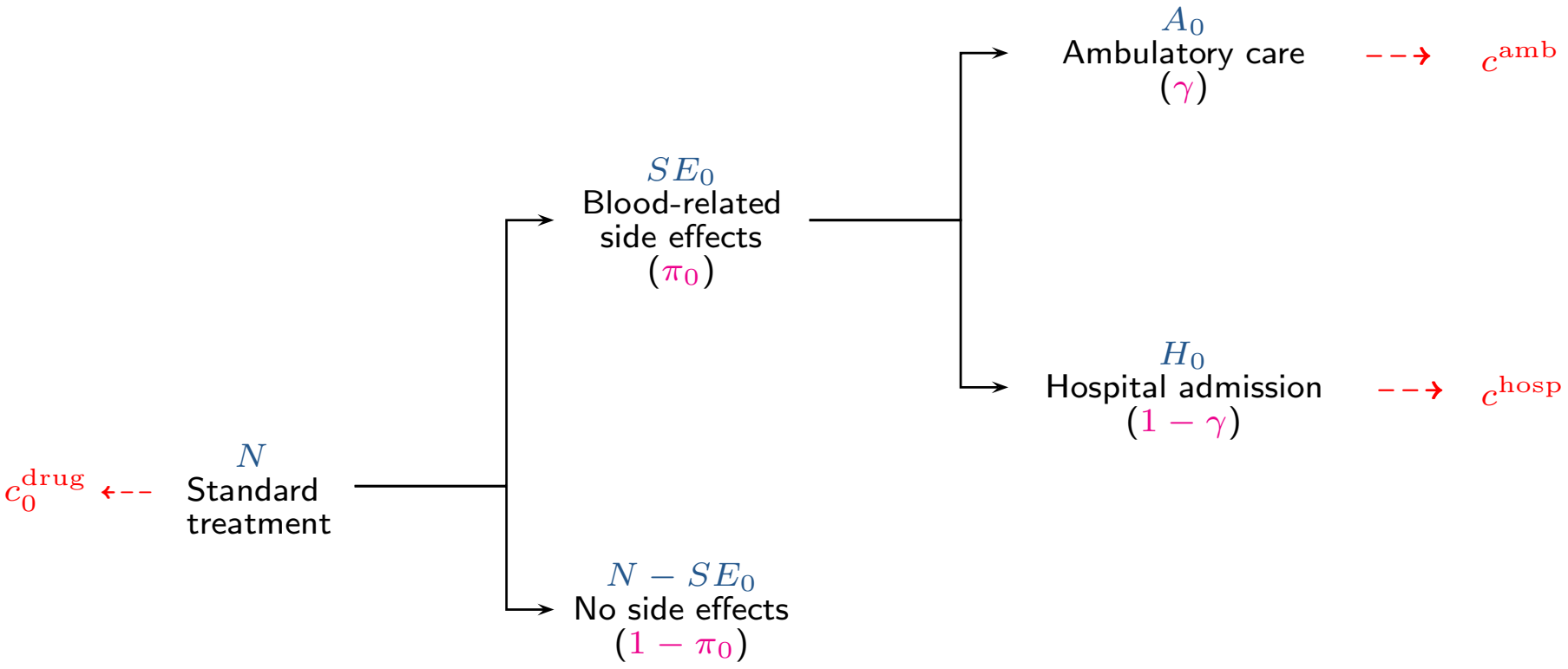
Example: Chemotherapy

$t = 0$: Old chemotherapy



Example: Chemotherapy

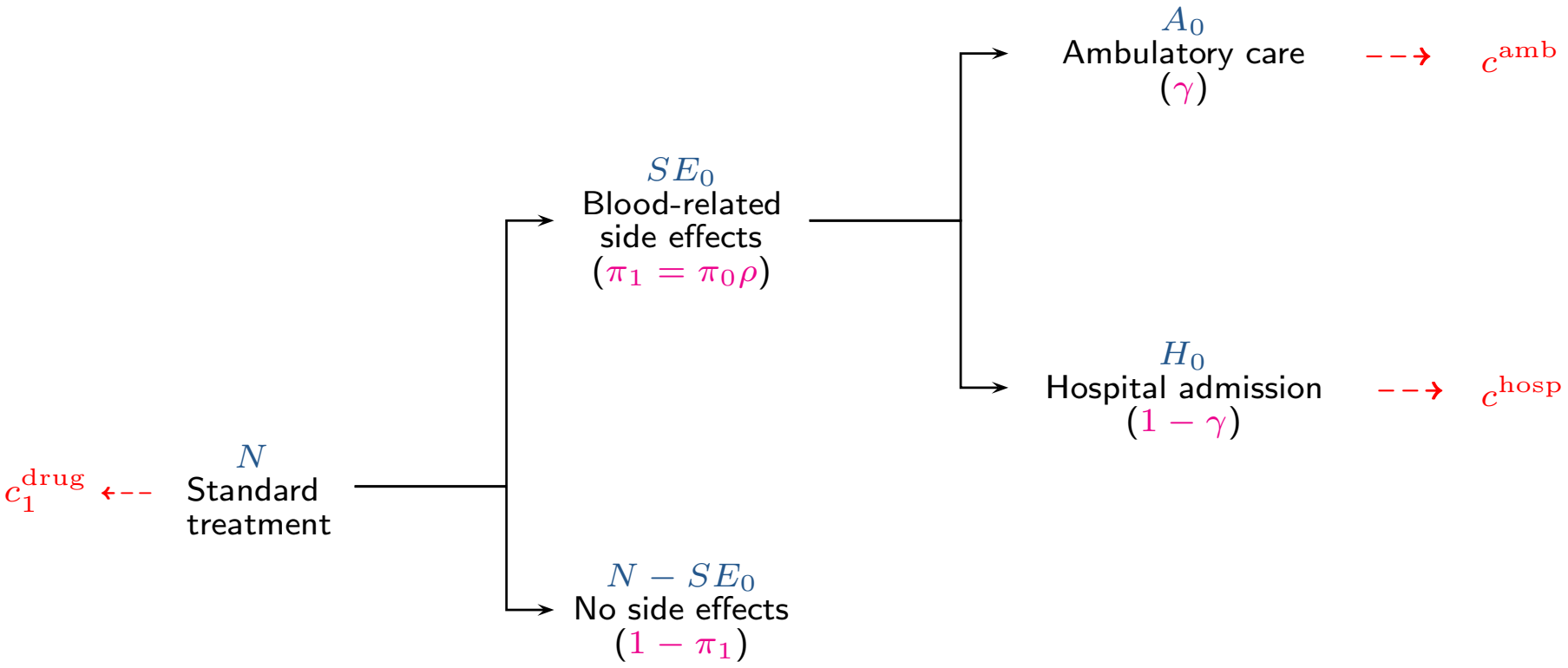
$t = 0$: Old chemotherapy



$$e_0 = N - SE_0 \qquad c_0 = Nc_0^{\text{drug}} + A_0c^{\text{amb}} + H_0c^{\text{hosp}}$$

Example: Chemotherapy

$t = 1$: New chemotherapy



$$e_1 = N - SE_1 \qquad c_1 = Nc_1^{\text{drug}} + A_1c^{\text{amb}} + H_1c^{\text{hosp}}$$

Expected Net Benefit

- Health economic decisions are based on the *utility* of a treatment, typically defined in terms of the monetary net benefit:

$$\text{nb}_t = ke_t - c_t$$

  where k is the willingness-to-pay.

- Uncertainty in this value is driven by e and c and an underlying parameter set $\boldsymbol{\theta}$

$$\boldsymbol{\theta} = (\pi_0, \gamma, \rho, SE_1, SE_2, A_1, A_1, H_1, H_2, c^{amb}, c^{hosp}, c_1^{drug}, c_2^{drug})$$

- To make decisions we maximise expected utility:

$$\mathcal{NB}_t = k\text{E}[e_t] - \text{E}[c_t]$$

- We typically wish to characterise the impact of parameter uncertainty using the known distribution utility

$$\text{NB}(\boldsymbol{\theta})_t = k\text{E}[e_t \mid \boldsymbol{\theta}] - \text{E}[c_t \mid \boldsymbol{\theta}]$$

  on the decision making process.

Value of Information

- Value of information methods can be used to summarise this parameter uncertainty
- A common summary is known as the Expected Value of Perfect Information

$$\text{EVPI} = \text{E}_{\boldsymbol{\theta}}\left[\max_t \{\text{NB}_t(\boldsymbol{\theta})\}\right] - \max_t \text{E}_{\boldsymbol{\theta}}\left[\text{NB}_t(\boldsymbol{\theta})\right]$$

- This gives an upper limit on future research costs
- Often we are concerned with research targeting a subset of parameters ϕ, e.g. $\phi = (\pi_1, \pi_2)$
- This is known as the Expected Value of *Partial* Perfect Information (EVPPI)

$$\text{EVPPI} = \text{E}_{\boldsymbol{\phi}}\left[\max_t \left\{\text{E}_{\boldsymbol{\psi}|\boldsymbol{\phi}}\left[\text{NB}_t(\boldsymbol{\theta})\right]\right\}\right] - \max_t \text{E}_{\boldsymbol{\psi},\boldsymbol{\phi}}\left[\text{NB}_t(\boldsymbol{\theta})\right]$$

where $\boldsymbol{\theta} = (\boldsymbol{\phi}, \boldsymbol{\psi})$

EVPPI as a regression problem

- Computational challenges have limited the applicability of EVPPI
- The calculation of the conditional expectation of the net benefit can be transformed into a regression problem

$$\text{NB}_t(\boldsymbol{\theta}) = \text{E}_{\boldsymbol{\psi}|\boldsymbol{\phi}}\left[\text{NB}_t(\boldsymbol{\theta})\right] + \epsilon$$

where $\epsilon \sim N(0, \sigma^2)$

- The conditional expectation is dependent on the value of $\boldsymbol{\phi}$

$$\text{NB}_t(\boldsymbol{\theta}) = g_t(\boldsymbol{\phi}) + \epsilon$$

- So to calculate the EVPPI we must find the functions $g_t(\boldsymbol{\phi})$

$$\widehat{\text{EVPPI}} = \frac{1}{S}\sum_{s=1}^{S} \max_t \hat{g}_t(\boldsymbol{\phi}_s) - \max_t \frac{1}{S}\sum_{s=1}^{S} \hat{g}_t(\boldsymbol{\phi}_s)$$

where S is the number of samples from the distribution of $\boldsymbol{\theta}$.

- Flexible, non-parametric regression methods should be used

Strong et al. (2014) [3]

Gaussian Process Regression

- Models the outputs as a multivariate normal dependent on some inputs ϕ
- Based on a mean function and a covariance function
  - Mean function based on the inputs, often linearly
  - Covariance function defines how correlated outputs are based on the inputs (often the distance between the inputs)
- These functions are given generic forms based on hyperparameters ζ
- We approximate these hyperparameters based on data
- MAP estimates are available but computationally costly

For example:

$$\begin{pmatrix} \text{NB}_t(\boldsymbol{\theta}_1) \\ \text{NB}_t(\boldsymbol{\theta}_2) \\ \vdots \\ \text{NB}_t(\boldsymbol{\theta}_S) \end{pmatrix} \sim \text{Normal}\left(\begin{pmatrix} 1 & \pi_1^1 & \pi_2^1 \\ 1 & \pi_1^2 & \pi_2^2 \\ \vdots & & \vdots \\ 1 & \pi_1^S & \pi_2^S \end{pmatrix} \boldsymbol{\beta}, \boldsymbol{\mathcal{C}}(\boldsymbol{\zeta}) + \sigma^2 \boldsymbol{I} \right)$$

INLA

- Integrated Nested Laplace Approximations (INLA) is a fast Bayesian inference method for Latent Gaussian Models.

$$y_i \mid \boldsymbol{\gamma}, \boldsymbol{\lambda} \sim \text{Dist}(h(\eta_i))$$

$$\eta_i = \alpha + \sum_{j=1}^{n_f} f_j(\gamma_{ji}) + \sum_{k=1}^{n_\beta} \beta_k \gamma_{ki} + \epsilon_i$$

$$\boldsymbol{\gamma}|\boldsymbol{\lambda} \sim N(\boldsymbol{\mu}(\boldsymbol{\lambda}), \boldsymbol{Q}^{-1}(\boldsymbol{\lambda}))$$

$$\boldsymbol{\lambda} \sim \pi(\boldsymbol{\lambda})$$

- $\boldsymbol{Q}(\boldsymbol{\lambda})$ must be sparse to allow for fast computation
- In order to use INLA, we must transform our Gaussian Process structure into a Latent Gaussian Field

Latent Gaussian Field

- We can rewrite our Gaussian process regression, with $\boldsymbol{H}$ as the design matrix, to mimic the Latent Gaussian Field structure:

$$\mathbf{NB}_t|\boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{\zeta} \sim N(\boldsymbol{H}\boldsymbol{\beta} + \boldsymbol{\omega}, \sigma^2\boldsymbol{I})$$

$$\eta_i = H_i\boldsymbol{\beta} + \omega_i$$

$$\begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\omega} \end{pmatrix} \sim N\left(\mathbf{0}, \begin{pmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\beta}} & 0 \\ 0 & \boldsymbol{Q}^{-1}(\boldsymbol{\zeta}) \end{pmatrix}\right)$$

$$\boldsymbol{\zeta} \sim \pi(\boldsymbol{\zeta})$$

- This is a Latent Gaussian Field if $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$ and $\boldsymbol{Q}(\boldsymbol{\zeta})$ are sparse matrices.
  - We assume that $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$ is known and sparse
  - $\boldsymbol{Q}(\boldsymbol{\zeta})$ is the covariance matrix which is not sparse but ideas developed in spatial statistics have allowed us to approximate this matrix by a sparse matrix

SPDE-INLA to calculate EVPPI

- INLA can be used in a spatial setting where the position of points has an impact on their respective values
- A Gaussian Process with a specific covariance function is the solution to a stochastic differential equation:

$$(\kappa^2 - \Delta)^{\frac{\alpha}{2}} \tau f(\phi) = \mathcal{W}(\phi)$$

  where Δ is the Laplcien and $\mathcal{W}(\phi)$ is Gaussian white noise.
- Therefore, approximating the solution of Stochastic Partial Differential Equations (SPDE) is equivalent to approximating our Matérn Gaussian Process
- Using the *finite element representation* we transform the estimation of ω into the estimation of a set of Gaussian weights with a sparse precision matrix.

Lindgren and Rue (2013) [2]

Projections

- This sparse precision matrix is only available in two dimensions
- The parameter set ϕ will often have more than two parameters
- Project from this higher dimensional space to 2 dimensions and then find the sparse precision matrix
- Use Principal Components Analysis as it preserves Euclidean distance
- The original values of ϕ are used to estimate $\boldsymbol{\beta}$

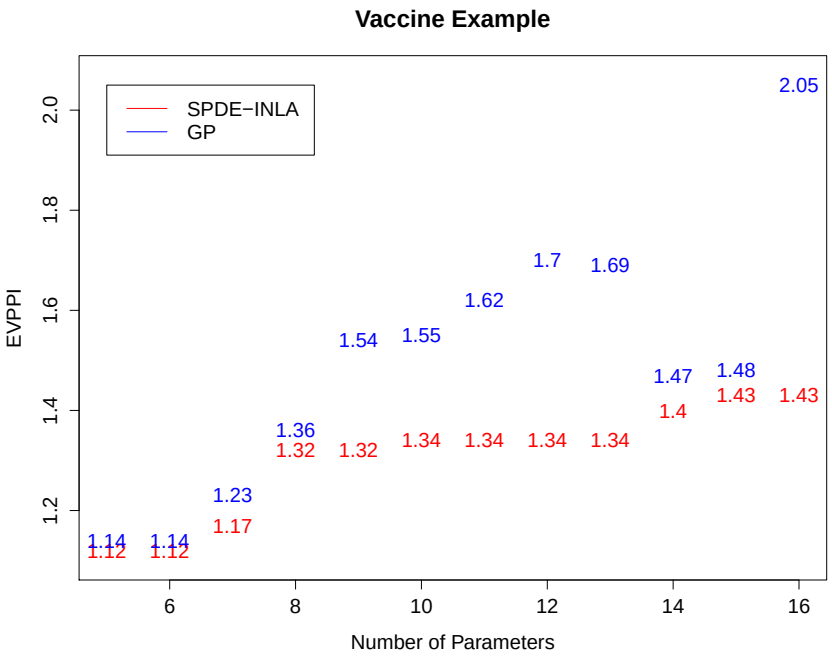
$$\mathbf{NB}_t|\boldsymbol{\omega}, \boldsymbol{\beta}, \boldsymbol{\zeta} \sim N(\boldsymbol{H\beta} + \boldsymbol{\omega}, \sigma^2 \boldsymbol{I})$$

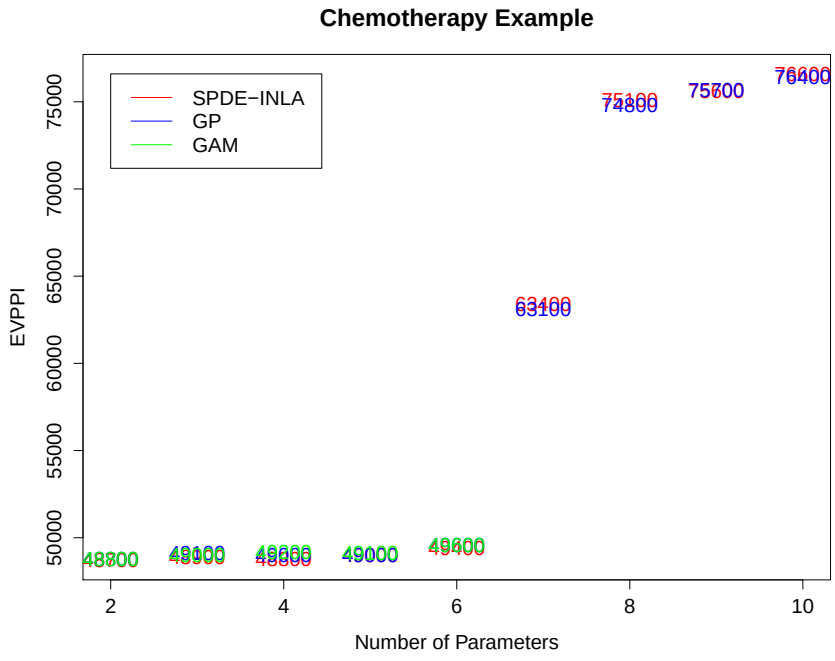
Heath et al. (2015) [1]

Computational Time

Number of important parameters	Computation Time			
	Vaccine Example		Chemotherapy	
	GP	SPDE-INLA	GP	SPDE-INLA
2	-	-	19	14
3	-	-	18	14
4	-	-	21	15
5	24	9	20	16
6	46	9	56	16
7	222	9	32	19
8	128	9	117	18
9	252	8	187	18
10	198	11	374	19
11	776	8	-	-
12	264	11	-	-
13	660	13	-	-
14	695	12	-	-
15	910	11	-	-
16	559	13	-	-

Accuracy


Vaccine Example
SPDE−INLA
GP
EVPPI
Number of Parameters
1.2
1.4
1.6
1.8
2.0
6
8
10
12
14
16
1.14
1.14
1.12
1.12
1.23
1.17
1.36
1.32
1.32
1.54
1.55
1.34
1.62
1.34
1.7
1.34
1.69
1.34
1.47
1.4
1.48
1.43
2.05
1.43


Chemotherapy Example
SPDE−INLA
GP
GAM
EVPPI
Number of Parameters
50000
55000
60000
65000
70000
75000
2
4
6
8
10
48800
49000
49600
63100
74800
75700
76400

Conclusion

- VoI methods are theoretically valid measures of decision uncertainty but their application has been hindered by the computational cost involved in calculating the EVPPI
- Strong et al. provide an efficient method to calculate the EVPPI but in some cases this is still expensive
- We have developed a method that calculates the EVPPI in around 10 seconds (for 1000 samples) irrespective of the complexity of the situation
- This methods draws on methods from spatial statistics and uses `R-INLA`
- Functions are available to allow practitioners to use this method easily and therefore calculate the EVPPI in all situations in around 10 seconds.

References

[1] A. Heath, I. Manolopoulou, and G. Baio. Efficient High-Dimensional Gaussian Process Regression to calculate the Expected Value of Partial Perfect Information in Health Economic Evaluations. *arXiv:1504.05436 [stat.AP]*, 2015.

[2] F. Lindgren and H. Rue. Bayesian spatial and spatiotemporal modelling with R-INLA. *Journal of Statistical Software*, 2013.

[3] Strong, M. and Oakley, J. and Brennan, A. Estimating Multiparameter Partial Expected Value of Perfect Information from a Probabilistic Sensitivity Analysis Sample: A Nonparametric Regression Approach. *Medical Decision Making*, 34(3):311–326, 2014.