Quick & clean: computationally efficient methods for Value of Information measures

Gianluca Baio

University College London
Department of Statistical Science

g.baio@ucl.ac.uk

(Joint work with Anna Heath and Ioanna Manolopoulou)
(Thanks to Mark Strong)

http://www.ucl.ac.uk/statistics/research/statistics-health-economics/
http://www.statistica.it/gianluca
https://github.com/giabaio

http://www.statistica.it/gianluca/project/voi/
https://www.convoi-group.org/

O'Bayes 2019
Objective Bayes Methodology Conference

University of Warwick, Monday 1 July 2019

O’Bayes?. . .

Outline

Outline

1. Value of Information
   - Basics

2. **EVPPI**
   - EVPPI as a (Gaussian Process) regression problem
   - Faster EVPPI (using INLA/SPDE)
   - Examples — research prioritisation

3. EVSI (I'll try really hard to keep time for this!...)
   - Ridiculously short introduction
   - Moment-matching
   - Examples — research prioritisation + study design

4. Conclusions

Outline

1. Value of Information
   - Basics
2. EVPPI
   - EVPPI as a (Gaussian Process) regression problem
   - Faster EVPPI (using INLA/SPDE)
   - Examples — research prioritisation
3. **EVSI (I'll try really hard to keep time for this!...)**
   - Ridiculously short introduction
   - Moment-matching
   - Examples — research prioritisation + study design
4. Conclusions

Outline

1. Value of Information
   - Basics
2. EVPPI
   - EVPPI as a (Gaussian Process) regression problem
   - Faster EVPPI (using INLA/SPDE)
   - Examples — research prioritisation
3. EVSI (I'll try really hard to keep time for this!...)
   - Ridiculously short introduction
   - Moment-matching
   - Examples — research prioritisation + study design
4. **Conclusions**

Objective: Combine costs & benefits of a given intervention (against at least one comparator) into a rational scheme for allocating resources

Health technology assessment (HTA)

Objective: Combine costs & benefits of a given intervention (against at least one comparator) into a rational scheme for allocating resources

Statistical model

- Estimates relevant **population** parameters $\boldsymbol{\theta}$
- Varies with the type of available data (& statistical approach!)

Health technology assessment (HTA)

Objective: Combine costs & benefits of a given intervention (against at least one comparator) into a rational scheme for allocating resources

$$\Delta_e = f_e(\boldsymbol\theta)$$
$$\Delta_c = f_c(\boldsymbol\theta)$$

- Estimates relevant **population** parameters $\boldsymbol\theta$
- Varies with the type of available data (& statistical approach!)

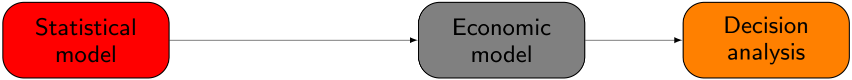
- Combines the parameters to obtain a population average measure for costs and clinical benefits
- Varies with the type of available data & statistical model used

Health technology assessment (HTA)

Objective: Combine costs & benefits of a given intervention (against at least one comparator) into a rational scheme for allocating resources

$$\Delta_e = f_e(\boldsymbol{\theta})$$
$$\Delta_c = f_c(\boldsymbol{\theta})$$

$$\text{ICER} = g(\Delta_e, \Delta_c)$$
$$\text{EIB} = h(\Delta_e, \Delta_c; k)$$
$$\ldots$$



Statistical model

- Estimates relevant **population** parameters $\boldsymbol{\theta}$
- Varies with the type of available data (& statistical approach!)

Economic model

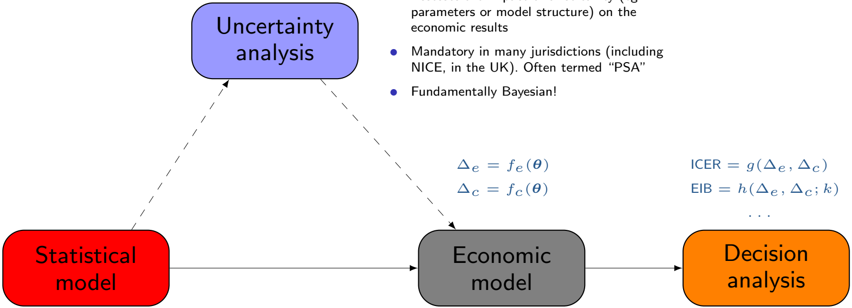
- Combines the parameters to obtain a population average measure for costs and clinical benefits
- Varies with the type of available data & statistical model used

Decision analysis

- Summarises the economic model by computing suitable measures of "cost-effectiveness"
- Dictates the best course of actions, given current evidence
- Standardised process

Health technology assessment (HTA)

Objective: Combine costs & benefits of a given intervention (against at least one comparator) into a rational scheme for allocating resources



- Assesses the impact of uncertainty (eg in parameters or model structure) on the economic results
- Mandatory in many jurisdictions (including NICE, in the UK). Often termed "PSA"
- Fundamentally Bayesian!

$$\Delta_e = f_e(\boldsymbol\theta)$$
$$\Delta_c = f_c(\boldsymbol\theta)$$

$$\text{ICER} = g(\Delta_e, \Delta_c)$$
$$\text{EIB} = h(\Delta_e, \Delta_c; k)$$
$$\ldots$$

- Estimates relevant **population** parameters $\boldsymbol\theta$
- Varies with the type of available data (& statistical approach!)

- Combines the parameters to obtain a population average measure for costs and clinical benefits
- Varies with the type of available data & statistical model used

- Summarises the economic model by computing suitable measures of "cost-effectiveness"
- Dictates the best course of actions, given current evidence
- Standardised process


PS4
Only On PlayStation.
PlayLink
KNOWLEDGE
is
POWER
3
REQUIRED

- **Example 1**: Intervention $t = 1$ is the most cost-effective, given current evidence
  - $\Pr(t = 1$ is cost-effective$) = 0.51$
  - If we get it wrong: Increase in costs = £3
    Decrease in effectiveness = 0.000001 QALYs
  - Large uncertainty/negligible consequences ⇒ **can afford uncertainty**
  - Best course of action: make a decision **now** (on the basis of current evidence)

Knowledge *is* power?... *(A tale of two stupid examples)*

- **Example 1**: Intervention $t = 1$ is the most cost-effective, given current evidence
  - $\Pr(t = 1 \text{ is cost-effective}) = 0.51$
  - If we get it wrong: Increase in costs = £3
    Decrease in effectiveness = 0.000001 QALYs
  - Large uncertainty/negligible consequences ⇒ **can afford uncertainty**
  - Best course of action: make a decision **now** (on the basis of current evidence)

- **Example 2**: Intervention $t = 1$ is the most cost-effective, given current evidence
  - $\Pr(t = 1 \text{ is cost-effective}) = 0.999$
  - If we get it wrong: Increase in costs = £1 000 000 000
    Decrease in effectiveness = 999999 QALYs
  - Tiny uncertainty/dire consequences ⇒ **probably should think about it...**
  - Best course of action: **gather more evidence** before making a decision

Value of Information (VoI): Basic idea

- A new study will provide new data
  - Reducing (ideally, *eliminating*) uncertainty in a subset of model parameters
- Update the cost-effectiveness model
  - If the optimal decision changes, gain in **monetary net benefit** (NB = utility) from using new optimal treatment
  - If optimal decision unchanged, no gain in NB
- **Expected** VOI is the average gain in NB

Value of Information (VoI): Basic idea and relevant measures

- A new study will provide new data
  - Reducing (ideally, *eliminating*) uncertainty in a subset of model parameters
- Update the cost-effectiveness model
  - If the optimal decision changes, gain in **monetary net benefit** (NB = utility) from using new optimal treatment
  - If optimal decision unchanged, no gain in NB
- **Expected** VOI is the average gain in NB

1. **Expected Value of Perfect Information** (EVPI)
   - Value of *completely resolving* uncertainty in *all* input parameters to decision model
   - Infinite-sized long-term follow-up trial measuring everything!
   - Gives an upper-bound on the value of new study — if EVPI is low, suggests we can make our decision based on existing information

Value of Information (VoI): Basic idea and relevant measures

- A new study will provide new data
  - Reducing (ideally, *eliminating*) uncertainty in a subset of model parameters
- Update the cost-effectiveness model
  - If the optimal decision changes, gain in **monetary net benefit** (NB = utility) from using new optimal treatment
  - If optimal decision unchanged, no gain in NB
- **Expected** VOI is the average gain in NB

1. **Expected Value of Perfect Information** (EVPI)
   - Value of *completely resolving* uncertainty in *all* input parameters to decision model
   - Infinite-sized long-term follow-up trial measuring everything!
   - Gives an upper-bound on the value of new study — if EVPI is low, suggests we can make our decision based on existing information
2. **Expected Value of Partial Perfect Information** (EVPPI)
   - Value of *completely resolving* uncertainty in *subset* of input parameters
   - Infinite-sized trial measuring relative effects on 1-year survival
   - Useful to identify which parameters responsible for decision uncertainty

Value of Information (VoI): Basic idea and relevant measures

- A new study will provide new data
  - Reducing (ideally, *eliminating*) uncertainty in a subset of model parameters
- Update the cost-effectiveness model
  - If the optimal decision changes, gain in **monetary net benefit** (NB = utility) from using new optimal treatment
  - If optimal decision unchanged, no gain in NB
- **Expected** VOI is the average gain in NB

1. **Expected Value of Perfect Information** (EVPI)
   - Value of *completely resolving* uncertainty in *all* input parameters to decision model
   - Infinite-sized long-term follow-up trial measuring everything!
   - Gives an upper-bound on the value of new study — if EVPI is low, suggests we can make our decision based on existing information
2. **Expected Value of Partial Perfect Information** (EVPPI)
   - Value of *completely resolving* uncertainty in *subset* of input parameters
   - Infinite-sized trial measuring relative effects on 1-year survival
   - Useful to identify which parameters responsible for decision uncertainty
3. **Expected Value of Sample Information** (EVSI)
   - Value of *reducing* uncertainty by conducting a finite-sized study of given design
   - Can compare the benefits and costs of different designs
   - Is the proposed study likely to be a good use of resources? What is the optimal design?

Value of Information (VoI): Basic idea and relevant measures

- A new study will provide new data
  - Reducing (ideally, *eliminating*) uncertainty in a subset of model parameters
- Update the cost-effectiveness model
  - If the optimal decision changes, gain in **monetary net benefit** (NB = utility) from using new optimal treatment
  - If optimal decision unchanged, no gain in NB
- **Expected** VOI is the average gain in NB

1. **Expected Value of Perfect Information** (EVPI)
   - Value of *completely resolving* uncertainty in *all* input parameters to decision model
   - Infinite-sized long-term follow-up trial measuring everything!
   - Gives an upper-bound on the value of new study — if EVPI is low, suggests we can make our decision based on existing information
2. **Expected Value of Partial Perfect Information** (EVPPI)
   - Value of *completely resolving* uncertainty in *subset* of input parameters
   - Infinite-sized trial measuring relative effects on 1-year survival
   - Useful to identify which parameters responsible for decision uncertainty
3. **Expected Value of Sample Information** (EVSI)
   - Value of *reducing* uncertainty by conducting a finite-sized study of given design
   - Can compare the benefits and costs of different designs
   - Is the proposed study likely to be a good use of resources? What is the optimal design?

- $\boldsymbol{\theta} =$ all the model parameters; can be split into two subsets
  - The "**parameters of interest**" ϕ, e.g. prevalence of a disease, HRQL measures, length of stay in hospital, ...
  - The "**remaining parameters**" ψ, e.g. cost of treatment with other established medications,
- We are interested in quantifying the value of gaining more information on ϕ, while leaving the current level of uncertainty on ψ unchanged

- $\boldsymbol\theta$ = all the model parameters; can be split into two subsets
  - The "**parameters of interest**" ϕ, e.g. prevalence of a disease, HRQL measures, length of stay in hospital, ...
  - The "**remaining parameters**" ψ, e.g. cost of treatment with other established medications,
- We are interested in quantifying the value of gaining more information on ϕ, while leaving the current level of uncertainty on ψ unchanged
- In formulæ:
  - First, consider the expected utility (EU) **if** we were able to learn ϕ but not ψ

$$\mathsf{E}_{\psi|\phi}\left[\mathsf{NB}_t(\boldsymbol\theta)\right]$$

- $\boldsymbol\theta$ = all the model parameters; can be split into two subsets
  - The "**parameters of interest**" ϕ, e.g. prevalence of a disease, HRQL measures, length of stay in hospital, ...
  - The "**remaining parameters**" ψ, e.g. cost of treatment with other established medications,
- We are interested in quantifying the value of gaining more information on ϕ, while leaving the current level of uncertainty on ψ unchanged
- In formulæ:
  - First, consider the expected utility ( ) **if** we were able to learn ϕ but not ψ
  - **If** we knew ϕ perfectly, best decision = the maximum of this EU

$$\max_t \mathrm{E}_{\psi\mid\phi}\left[\mathrm{NB}_t(\boldsymbol\theta)\right]$$

- $\boldsymbol{\theta}$ = all the model parameters; can be split into two subsets
  - The "**parameters of interest**" ϕ, e.g. prevalence of a disease, HRQL measures, length of stay in hospital, ...
  - The "**remaining parameters**" ψ, e.g. cost of treatment with other established medications,
- We are interested in quantifying the value of gaining more information on ϕ, while leaving the current level of uncertainty on ψ unchanged
- In formulæ:
  - First, consider the expected utility ( ) **if** we were able to learn ϕ but not ψ
  - **If** we knew ϕ perfectly, best decision = the maximum of this EU
  - Of course we cannot learn ϕ **perfectly**, so take the expected value

$$\mathsf{E}_{\boldsymbol{\phi}}\left[\max_t \mathsf{E}_{\psi|\phi}\left[\mathsf{NB}_t(\boldsymbol{\theta})\right]\right]$$

- $\boldsymbol\theta$ = all the model parameters; can be split into two subsets
  - The "**parameters of interest**" ϕ, e.g. prevalence of a disease, HRQL measures, length of stay in hospital, ...
  - The "**remaining parameters**" ψ, e.g. cost of treatment with other established medications,
- We are interested in quantifying the value of gaining more information on ϕ, while leaving the current level of uncertainty on ψ unchanged
- In formulæ:
  - First, consider the expected utility ( ) **if** we were able to learn ϕ but not ψ
  - **If** we knew ϕ perfectly, best decision = the maximum of this EU
  - Of course we cannot learn ϕ **perfectly**, so take the expected value
  - And compare this with the **maximum expected utility overall**

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

Summarising PSA + Research priority: Expected Value of **Partial** Perfect Information

- $\boldsymbol\theta$ = all the model parameters; can be split into two subsets
  - The "**parameters of interest**" ϕ, e.g. prevalence of a disease, HRQL measures, length of stay in hospital, ...
  - The "**remaining parameters**" ψ, e.g. cost of treatment with other established medications,
- We are interested in quantifying the value of gaining more information on ϕ, while leaving the current level of uncertainty on ψ unchanged
- In formulæ:
  - First, consider the expected utility ( ) **if** we were able to learn ϕ but not ψ
  - **If** we knew ϕ perfectly, best decision = the maximum of this EU
  - Of course we cannot learn ϕ **perfectly**, so take the expected value
  - And compare this with the **maximum expected utility overall**
  - This is the EVPPI!

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

Summarising PSA + Research priority: Expected Value of **Partial** Perfect Information

- $\boldsymbol\theta$ = all the model parameters; can be split into two subsets
  - The "**parameters of interest**" ϕ, e.g. prevalence of a disease, HRQL measures, length of stay in hospital, ...
  - The "**remaining parameters**" ψ, e.g. cost of treatment with other established medications,
- We are interested in quantifying the value of gaining more information on ϕ, while leaving the current level of uncertainty on ψ unchanged
- In formulæ:
  - First, consider the expected utility ( ) **if** we were able to learn ϕ but not ψ
  - **If** we knew ϕ perfectly, best decision = the maximum of this EU
  - Of course we cannot learn ϕ **perfectly**, so take the expected value
  - And compare this with the **maximum expected utility overall**
  - This is the EVPPI!

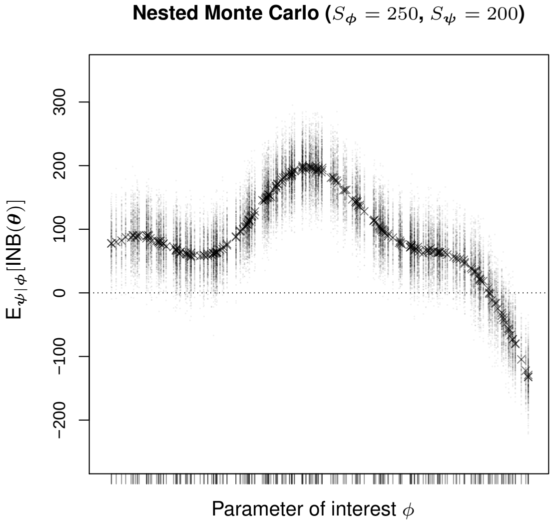
$$\text{EVPPI} = \text{E}_\phi \left[\max_t \text{E}_{\boldsymbol\psi|\boldsymbol\phi} \left[\text{NB}_t(\boldsymbol\theta)\right] \right] - \max_t \text{E}_{\boldsymbol\theta}\left[\text{NB}_t(\boldsymbol\theta)\right]$$

- **That's** the difficult part!
  - Can do nested Monte Carlo, but takes forever to get accurate results
  - **Recent methods** based on **Gaussian Process regression** very efficient & quick!

Strong et al *Medical Decision Making*. 2014; **34(3)**: 311-26.
Heath et al *Statistics in Medicine*. 2016; **35(23)**: 4264-4280.

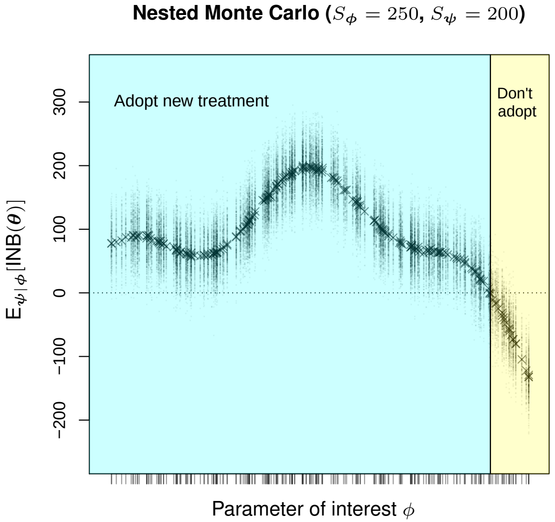
http://savi.shef.ac.uk/SAVI/
https://egon.stats.ucl.ac.uk/projects/EVSI/

Assuming only two interventions, can consider $\text{INB}(\boldsymbol{\theta}) = \text{NB}_1(\boldsymbol{\theta}) - \text{NB}_0(\boldsymbol{\theta})$



Thanks to Mark Strong (slide stolen from "Summer School in Bayesian methods in health economics")
www.statistica.it/gianluca/teaching/summer-school

Assuming only two interventions, can consider $\text{INB}(\boldsymbol{\theta}) = \text{NB}_1(\boldsymbol{\theta}) - \text{NB}_0(\boldsymbol{\theta})$



Thanks to Mark Strong (slide stolen from "Summer School in Bayesian methods in health economics")
www.statistica.it/gianluca/teaching/summer-school

- Instead of Nested Monte Carlo, can model as a **regression** problem

$$\begin{aligned} \text{NB}_t(\boldsymbol\theta) &= \text{E}_{\boldsymbol\psi|\boldsymbol\phi}\left[\text{NB}_t(\boldsymbol\theta)\right] + \varepsilon, \qquad \text{with } \varepsilon \sim \text{Normal}(0,\sigma^2_\varepsilon) \\ &= g_t(\boldsymbol\phi) + \varepsilon \end{aligned}$$

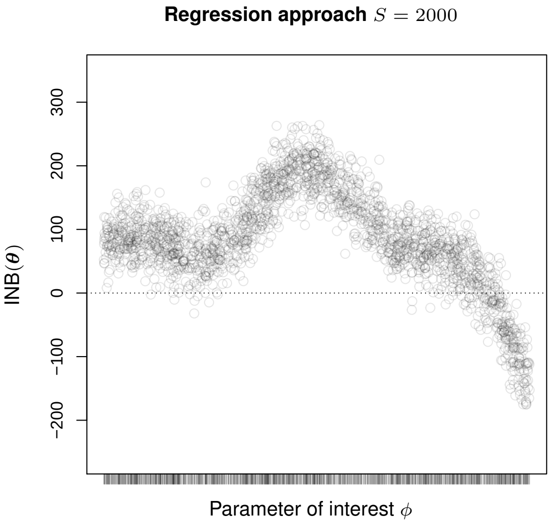
"Data": **simulations** for $\text{NB}_t(\boldsymbol\theta)$ as "response"
simulations for $\boldsymbol\phi$ as "covariates"

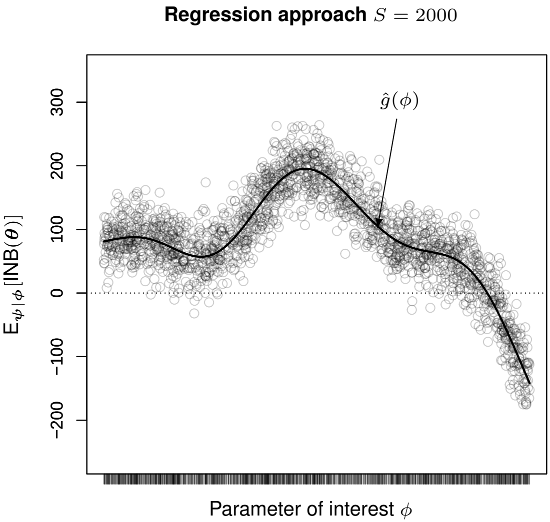
- **NB**: Only need S data points (= PSA simulations), instead of $S_{\boldsymbol\phi} \times S_{\boldsymbol\psi}$!

π_0	ρ	β_0	...	σ	η	γ	$\text{NB}_0(\boldsymbol\theta)$	$\text{NB}_1(\boldsymbol\theta)$
0.365	0.076	0.243	...	0.622	0.001	0.162	19 214 751	19 647 706
0.421	0.024	0.115	...	0.519	0.010	0.134	17 165 526	17 163 407
0.125	0.017	0.420	...	0.482	0.007	0.149	18 710 928	16 458 433
0.117	0.073	0.419	...	0.317	0.003	0.120	16 991 321	18 497 648
0.481	0.008	0.176	...	0.497	0.004	0.191	19 772 898	18 662 329
0.163	0.127	0.227	...	0.613	0.083	0.004	17 106 136	18 983 331
	...		...		...		...	
0.354	0.067	0.318	...	0.519	0.063	0.117	18 043 921	16 470 805

"covariates" "response" "response"

Strong et al *Medical Decision Making*. 2014; **34(3)**: 311-26


Regression approach $S = 2000$
INB($\boldsymbol{\theta}$)
300
200
100
0
−100
−200
Parameter of interest ϕ


Regression approach $S = 2000$
$\hat{g}(\phi)$
300
200
100
0
−100
−200
$E_{\psi|\phi}[\text{INB}(\boldsymbol{\theta})]$
Parameter of interest ϕ

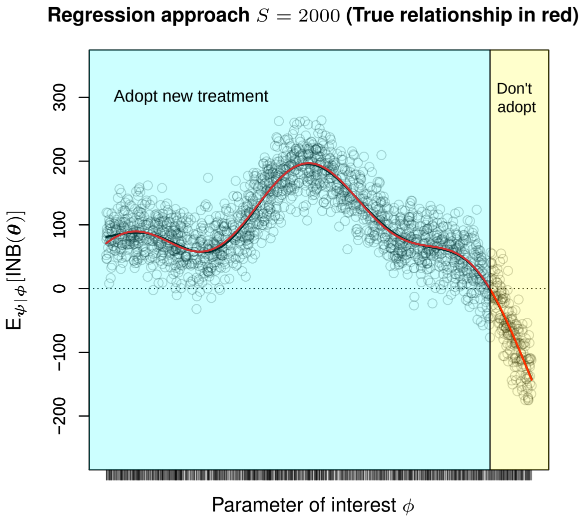
- Instead of Nested Monte Carlo, can model as a **regression** problem

$$\begin{aligned} \text{NB}_t(\boldsymbol\theta) &= \text{E}_{\boldsymbol\psi|\boldsymbol\phi}\left[\text{NB}_t(\boldsymbol\theta)\right] + \varepsilon, \qquad \text{with } \varepsilon \sim \text{Normal}(0, \sigma^2_\varepsilon) \\ &= g_t(\boldsymbol\phi) + \varepsilon \end{aligned}$$

"Data": **simulations** for $\text{NB}_t(\boldsymbol\theta)$ as "response"
simulations for $\boldsymbol\phi$ as "covariates"

- Once the functions $g_t(\boldsymbol\phi)$ are estimated, then can approximate

$$\begin{aligned} \text{EVPPI} &= \text{E}_{\boldsymbol\phi}\left[\max_t \text{E}_{\boldsymbol\psi|\boldsymbol\phi}\left[\text{NB}_t(\boldsymbol\theta)\right]\right] - \max_t \text{E}_{\boldsymbol\theta}\left[\text{NB}_t(\boldsymbol\theta)\right] \\ &\approx \frac{1}{S}\sum_{s=1}^S \max_t \hat{g}_t(\boldsymbol\phi_s) - \max_t \frac{1}{S}\sum_{s=1}^S \hat{g}_t(\boldsymbol\phi_s) \end{aligned}$$


Regression approach $S = 2000$ (True relationship in red)
Adopt new treatment
Don't adopt
300
200
100
0
−100
−200
$E_{\psi|\phi}[INB(\boldsymbol{\theta})]$
Parameter of interest ϕ

- Instead of Nested Monte Carlo, can model as a **regression** problem

$$\begin{aligned} \mathsf{NB}_t(\boldsymbol\theta) &= \mathsf{E}_{\boldsymbol\psi|\boldsymbol\phi}\left[\mathsf{NB}_t(\boldsymbol\theta)\right] + \varepsilon, \qquad \text{with } \varepsilon \sim \text{Normal}(0,\sigma^2_\varepsilon) \\ &= g_t(\boldsymbol\phi) + \varepsilon \end{aligned}$$

"Data": **simulations** for $\mathsf{NB}_t(\boldsymbol\theta)$ as "response"
simulations for $\boldsymbol\phi$ as "covariates"

- Once the functions $g_t(\boldsymbol\phi)$ are estimated, then can approximate

$$\begin{aligned} \text{EVPPI} &= \mathsf{E}_{\boldsymbol\phi}\left[\max_t \mathsf{E}_{\boldsymbol\psi|\boldsymbol\phi}\left[\mathsf{NB}_t(\boldsymbol\theta)\right]\right] - \max_t \mathsf{E}_{\boldsymbol\theta}\left[\mathsf{NB}_t(\boldsymbol\theta)\right] \\ &\approx \frac{1}{S}\sum_{s=1}^{S}\max_t \hat{g}_t(\boldsymbol\phi_s) - \max_t \frac{1}{S}\sum_{s=1}^{S}\hat{g}_t(\boldsymbol\phi_s) \end{aligned}$$

- **NB**: $g_t(\boldsymbol\phi)$ can be complex, so need to use **flexible** regression methods
  - **GAMs**: $g_t(\boldsymbol\phi) = \sum_{q=1}^{Q_{\boldsymbol\phi}} h_t(\phi_{sq})$ $\qquad h_t(\cdot) =$ smooth functions (cubic polynomials)
    very fast, but only work if number of important parameters $Q_{\boldsymbol\phi} \leq 5$ (interactions increase model size exponentially!)
  - If $P > 5$, can use **Gaussian Process** regression

EVPPI via **GP regression**

Model

$$\begin{pmatrix} \text{NB}_t(\boldsymbol\theta_1) \\ \text{NB}_t(\boldsymbol\theta_2) \\ \vdots \\ \text{NB}_t(\boldsymbol\theta_S) \end{pmatrix} := \textbf{NB}_t \sim \text{Normal}(\boldsymbol{H\beta}, \boldsymbol{\mathcal{C}}_{\text{Exp}} + \sigma^2_\varepsilon \boldsymbol I)$$

$$\boldsymbol H = \begin{pmatrix} 1 & \phi_{11} & \cdots & \phi_{1P} \\ 1 & \phi_{21} & \cdots & \phi_{2P} \\ \vdots & & \ddots & \\ 1 & \phi_{S1} & \cdots & \phi_{SP} \end{pmatrix} \quad \text{and} \quad \mathcal{C}_{\text{Exp}}(r,s) = \sigma^2 \exp\left[\sum_{p=1}^P \left(\frac{\phi_{rp}-\phi_{sp}}{\delta_p}\right)^2\right]$$

- Parameters: $\boldsymbol\beta$, $\boldsymbol\delta$, σ^2, σ^2_ε
- Very flexible structure — good approximation level
- Can use conjugate priors + numerical optimisation
- **But**: can still be very slow — computational cost in the order of S^3 (involves inversion of a dense covariance matrix)

EVPPI via GP regression — **but faster**

Heath et al *Stats in Med.* 2016; **35(23)**: 4264-4280

EVPPI via GP regression — **but faster**

1. Build from ideas in **spatial statistics** and use a Matérn covariance function

$$\mathcal{C}_{\rm M}(r,s) = \frac{\sigma^2}{\Gamma(\nu)2^{\nu-1}}(\kappa\|\boldsymbol\phi_r - \boldsymbol\phi_s\|)^\nu \mathsf{K}_\nu(\kappa\|\boldsymbol\phi_r - \boldsymbol\phi_s\|)$$

   - Fewer parameters, but still implies a dense covariance matrix
   - **But**: can use efficient algorithms to solve **Stochastic Partial Differential Equations** (SPDE) to approximate it — with computational cost $\propto S^{3/2}$!

Heath et al *Stats in Med.* 2016; **35(23)**: 4264-4280; Lindgren et al *JRSS/B*. 2011; **73(4)**: 423-498

EVPPI via GP regression — **but faster**

1. Build from ideas in **spatial statistics** and use a Matérn covariance function

$$\mathcal{C}_{\rm M}(r,s) = \frac{\sigma^2}{\Gamma(\nu)2^{\nu-1}}(\kappa\|\boldsymbol\phi_r - \boldsymbol\phi_s\|)^\nu \mathsf{K}_\nu(\kappa\|\boldsymbol\phi_r - \boldsymbol\phi_s\|)$$

   - Fewer parameters, but still implies a dense covariance matrix
   - **But**: can use efficient algorithms to solve **Stochastic Partial Differential Equations** (SPDE) to approximate it — with computational cost $\propto S^{3/2}$!

2. Re-formulate the model as

$$\begin{aligned}\textbf{NB}_t &\sim \text{Normal}(\boldsymbol{H\beta}, \boldsymbol{\mathcal{C}}_{\rm M} + \sigma^2_\varepsilon \boldsymbol I) \\ &\sim \text{Normal}(\boldsymbol{H\beta} + f(\boldsymbol\omega), \sigma^2_\varepsilon \boldsymbol I)\end{aligned}$$

   - $f(\boldsymbol\omega)$ are a set of "spatially structured" effects, with $\boldsymbol\omega \sim \text{Normal}\left(0, \boldsymbol Q^{-1}(\xi)\right)$
   - $\boldsymbol Q(\xi)$ is a **sparse** precision matrix determined by the SPDE solution
   - **Project** the $P-$dimensional information in $\boldsymbol\phi$ to a $d-$dimensional space ($d << P$, ideally $d=2$) — can use "Principal Fitted Components"

Heath et al *Stats in Med.* 2016; **35(23)**: 4264-4280; Lindgren et al *JRSS/B.* 2011; **73(4)**: 423-498

EVPPI via GP regression — **but faster**

1. Build from ideas in **spatial statistics** and use a Matérn covariance function

$$\mathcal{C}_{\rm M}(r,s) = \frac{\sigma^2}{\Gamma(\nu)2^{\nu-1}}(\kappa\|\boldsymbol\phi_r - \boldsymbol\phi_s\|)^\nu \mathsf{K}_\nu(\kappa\|\boldsymbol\phi_r - \boldsymbol\phi_s\|)$$

   - Fewer parameters, but still implies a dense covariance matrix
   - **But**: can use efficient algorithms to solve **Stochastic Partial Differential Equations** (SPDE) to approximate it — with computational cost $\propto S^{3/2}$!

2. Re-formulate the model as

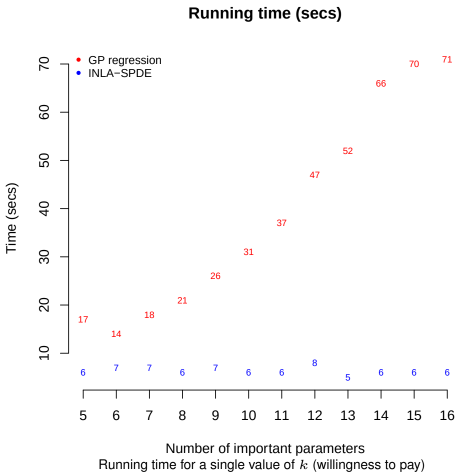
$$\begin{aligned}\textbf{NB}_t &\sim \text{Normal}(\boldsymbol H\boldsymbol\beta, \boldsymbol{\mathcal{C}}_{\rm M} + \sigma^2_\varepsilon \boldsymbol I)\\ &\sim \text{Normal}(\boldsymbol H\boldsymbol\beta + f(\boldsymbol\omega), \sigma^2_\varepsilon \boldsymbol I)\end{aligned}$$

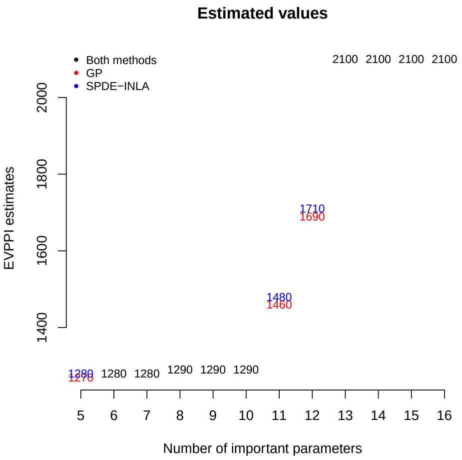
   - $f(\boldsymbol\omega)$ are a set of "spatially structured" effects, with $\boldsymbol\omega \sim \text{Normal}\left(0, \boldsymbol Q^{-1}(\xi)\right)$
   - $\boldsymbol Q(\xi)$ is a **sparse** precision matrix determined by the SPDE solution
   - **Project** the $P-$dimensional information in $\boldsymbol\phi$ to a $d-$dimensional space ($d << P$, ideally $d=2$) — can use "Principal Fitted Components"

3. Crucially, if we set a sparse Gaussian prior on $\boldsymbol\beta$, this is a Latent Gaussian model $\Rightarrow$ can estimate using super-fast **Integrated Nested Laplace Approximation** (INLA)

NB: Both GP-based methods implemented in the `R` package **BCEA** (Bayesian Cost-Effectiveness Analysis)
`http://www.statistica.it/gianluca/BCEA` `https://github.com/giabaio/BCEA`

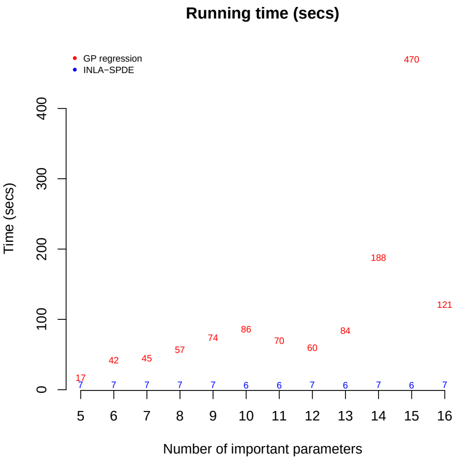
Heath et al *Stats in Med.* 2016; **35(23)**: 4264-4280; Lindgren et al *JRSS/B.* 2011; **73(4)**: 423-498; Rue et al *JRSS/B.* 2009; **71**: 319-392



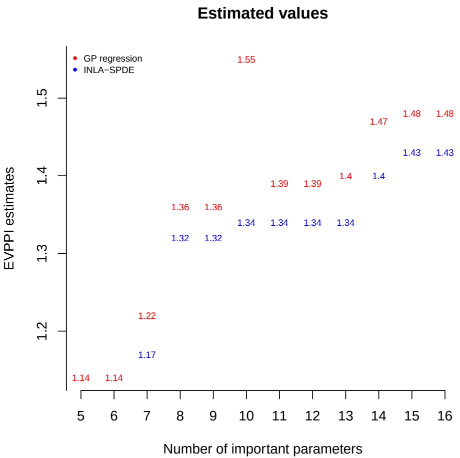


- Fictional decision tree model with correlated parameters
- 2 treatment options and overall 19 parameters
- Parameters simulated from multivariate Normal distribution, so can compute exact EVPPI

Heath et al *Statistics in Medicine*. 2016; **35(23)**: 4264-4280



Running time for a single value of k (willingness to pay)

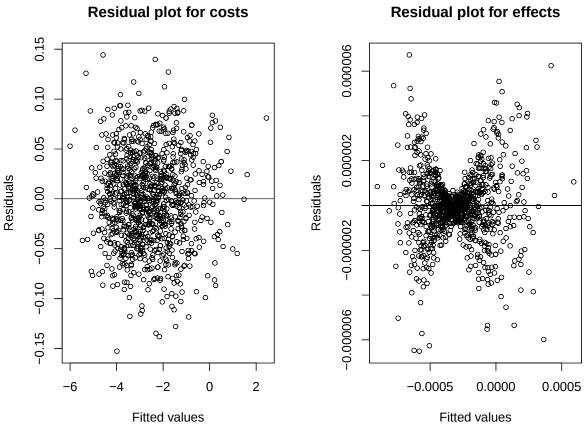


- Cost-effectiveness model for influenza vaccine based on evidence synthesis
- 2 treatment options and overall 63 parameters
- Model not available in closed form (needs MCMC simulations)

Heath et al *Statistics in Medicine*. 2016; **35(23)**: 4264-4280

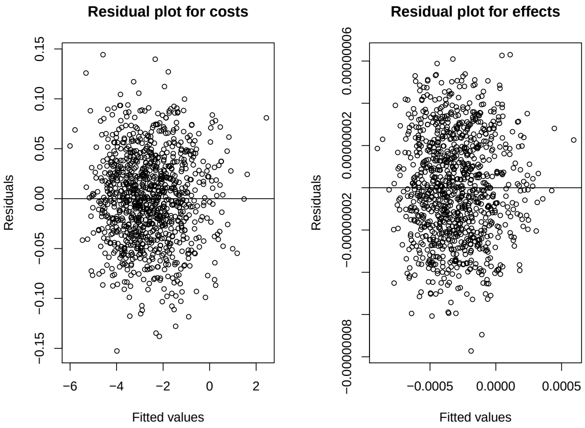
Breast cancer screening (Welton et al. 2008. *JRSS/A*)

- Multi-decision model developed for the UK setting, with 4 interventions
- Complex evidence synthesis for 6 parameters — highly structured!



- Can relatively easily modify the basic structure of the model, e.g. include interaction terms to make $\boldsymbol{H}\boldsymbol{\beta}$ non-linear

$$\beta_0 + \beta_1\phi_{1s} + \beta_2\phi_{2s} + \beta_3\phi_{3s} + \beta_4\phi_{1s}\phi_{2s} + \beta_5\phi_{1s}\phi_{3s} + \beta_6\phi_{2s}\phi_{3s}$$



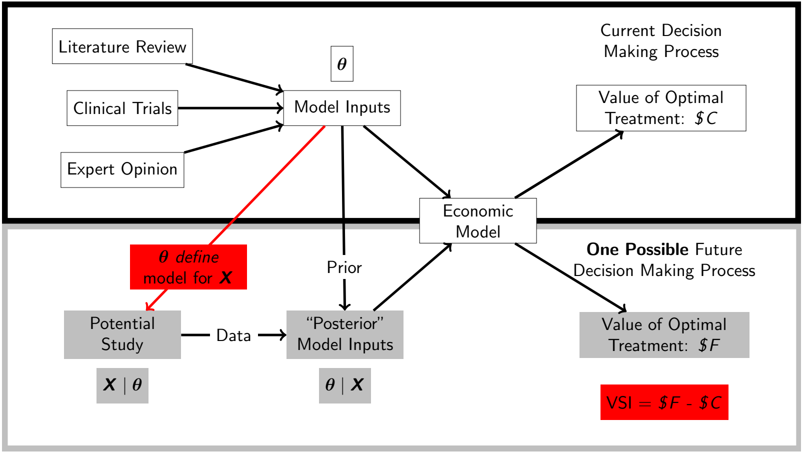
Baio et al (2017). *Bayesian Cost-Effectiveness Analysis with the R package BCEA*

Value of Information (VoI): Basic idea and relevant measures

- A new study will provide new data
  - Reducing (ideally, *eliminating*) uncertainty in a subset of model parameters
- Update the cost-effectiveness model
  - If the optimal decision changes, gain in **monetary net benefit** (NB = utility) from using new optimal treatment
  - If optimal decision unchanged, no gain in NB
- **Expected** VOI is the average gain in NB

1. **Expected Value of Perfect Information** (EVPI)
   - Value of *completely resolving* uncertainty in *all* input parameters to decision model
   - Infinite-sized long-term follow-up trial measuring everything!
   - Gives an upper-bound on the value of new study — if EVPI is low, suggests we can make our decision based on existing information
2. **Expected Value of Partial Perfect Information** (EVPPI)
   - Value of *completely resolving* uncertainty in *subset* of input parameters
   - Infinite-sized trial measuring relative effects on 1-year survival
   - Useful to identify which parameters responsible for decision uncertainty
3. **Expected Value of Sample Information** (EVSI)
   - Value of *reducing* uncertainty by conducting a finite-sized study of given design
   - Can compare the benefits and costs of different designs
   - Is the proposed study likely to be a good use of resources? What is the optimal design?

- EVSI measures the value of reducing uncertainty by running a study of a given design
- Can compare the benefits and costs of a study with given design
  - To see if a proposed study likely to be a good use of resources
  - To find the optimal study design



- EVSI measures the value of reducing uncertainty by running a study of a given design
- Can compare the benefits and costs of a study with given design
  - To see if a proposed study likely to be a good use of resources
  - To find the optimal study design

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

Posterior given data $\boldsymbol{X}$

Prior predictive distribution (pre-posterior)

Value of decision based on **sample** information (for a given study design)

Value of decision based on **current** information

- Computationally complex
  - Requires specific knowledge of the model for (future/hypothetical) data collection
  - Again, recent methods have improved efficiency (Heath et al. 2019; https://arxiv.org/abs/1905.12013)
- Can be used to drive design of new study (eg sample size calculations)

Objective: Estimate the distribution $p(\mu^{\boldsymbol X})$ with $\mu^{\boldsymbol X} = \text{E}_{\boldsymbol\theta\mid\boldsymbol X}[\text{INB}(\boldsymbol\theta)]$

- That's the hard part to estimate the EVSI

Objective: Estimate the distribution $p(\mu^{\boldsymbol{X}})$ with $\mu^{\boldsymbol{X}} = \mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}[\text{INB}(\boldsymbol{\theta})]$

- That's the hard part to estimate the EVSI

We know that

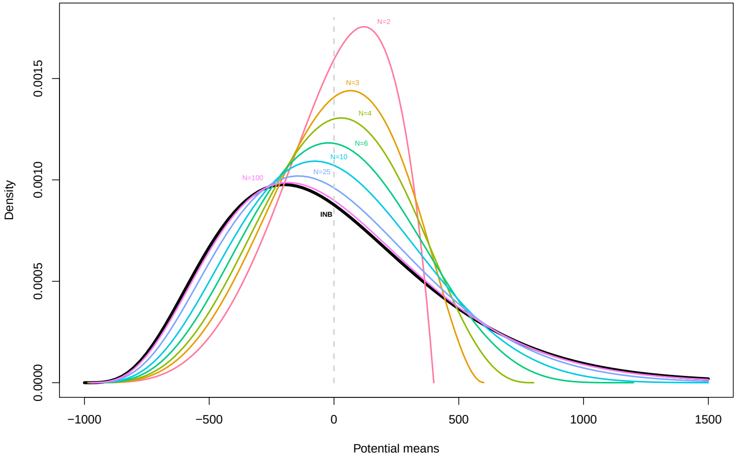
1. $\mathsf{E}_{\boldsymbol{X}}\left[\mu^{\boldsymbol{X}}\right] = \mathsf{E}_{\boldsymbol{X}}\left[\mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\text{INB}(\boldsymbol{\theta})\right]\right] = \mathsf{E}_{\boldsymbol{\theta}}\left[\text{INB}(\boldsymbol{\theta})\right]$
2. $\text{Var}_{\boldsymbol{X}}\left[\mu^{\boldsymbol{X}}\right] = \underbrace{\text{Var}_{\boldsymbol{\theta}}\left[\text{INB}(\boldsymbol{\theta})\right]}_{\text{PSA variance for INB}(\boldsymbol{\theta})} - \underbrace{\mathsf{E}_{\boldsymbol{X}}\left[\text{Var}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\text{INB}(\boldsymbol{\theta})\right]\right]}_{\text{Posterior variance for INB}(\boldsymbol{\theta})}$

Objective: Estimate the distribution $p(\mu^{\boldsymbol{X}})$ with $\mu^{\boldsymbol{X}} = \mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}[\mathsf{INB}(\boldsymbol{\theta})]$

- That's the hard part to estimate the EVSI

We know that

1. $\mathsf{E}_{\boldsymbol{X}}\left[\mu^{\boldsymbol{X}}\right] = \mathsf{E}_{\boldsymbol{X}}\left[\mathsf{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\mathsf{INB}(\boldsymbol{\theta})\right]\right] = \mathsf{E}_{\boldsymbol{\theta}}\left[\mathsf{INB}(\boldsymbol{\theta})\right]$
2. $\mathsf{Var}_{\boldsymbol{X}}\left[\mu^{\boldsymbol{X}}\right] = \underbrace{\mathsf{Var}_{\boldsymbol{\theta}}\left[\mathsf{INB}(\boldsymbol{\theta})\right]}_{\text{PSA variance for INB}(\boldsymbol{\theta})} - \underbrace{\mathsf{E}_{\boldsymbol{X}}\left[\mathsf{Var}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\mathsf{INB}(\boldsymbol{\theta})\right]\right]}_{\text{Posterior variance for INB}(\boldsymbol{\theta})}$
3. As $n \to \infty$, $p(\mu^{\boldsymbol{X}})$ is "similar" to the PSA distribution of INB$(\boldsymbol{\theta})$



- If $n \to \infty$, we will have learned the value of θ perfectly
  - But we haven't seen the new (infinite sized) data yet — so current uncertainty described by the prior
- If $n \to 0$, there are fewer values we can learn from future data
  - So pre-posterior more concentrated away from the prior

Objective: Estimate the distribution $p(\mu^{\boldsymbol{X}})$ with $\mu^{\boldsymbol{X}} = \text{E}_{\boldsymbol{\theta}|\boldsymbol{X}}[\text{INB}(\boldsymbol{\theta})]$

- That's the hard part to estimate the EVSI

We know that

1. $\text{E}_{\boldsymbol{X}}\left[\mu^{\boldsymbol{X}}\right] = \text{E}_{\boldsymbol{X}}\left[\text{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\text{INB}(\boldsymbol{\theta})\right]\right] = \text{E}_{\boldsymbol{\theta}}\left[\text{INB}(\boldsymbol{\theta})\right]$
2. $\text{Var}_{\boldsymbol{X}}\left[\mu^{\boldsymbol{X}}\right] = \underbrace{\text{Var}_{\boldsymbol{\theta}}\left[\text{INB}(\boldsymbol{\theta})\right]}_{\text{PSA variance for INB}(\boldsymbol{\theta})} - \underbrace{\text{E}_{\boldsymbol{X}}\left[\text{Var}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\text{INB}(\boldsymbol{\theta})\right]\right]}_{\text{Posterior variance for INB}(\boldsymbol{\theta})}$
3. As $n \to \infty$, $p(\mu^{\boldsymbol{X}})$ is "similar" to the PSA distribution of INB($\boldsymbol{\theta}$)

Idea: can approximate the unknown distribution $p(\mu^{\boldsymbol{X}})$ by rescaling the PSA distribution for INB($\boldsymbol{\theta}$), **moment-matching** it to the mean and variance defined above

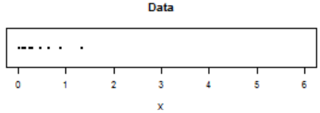
- All we need is to estimate the **expected posterior variance**...
- Can do this efficiently by only using $Q \approx 30 \text{ to } 50 << S$ PSA simulations!

Heath et al. 2017. *Medical Decision Making*. **38(2)**: 163-173

1. Select $q = 1, \ldots, Q$ values out of the S PSA simulations for $\boldsymbol{\theta}$

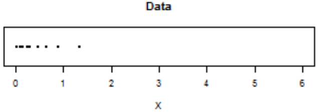
1. Select $q = 1, \ldots, Q$ values out of the S PSA simulations for $\boldsymbol\theta$

2. Simulate data $\boldsymbol X_q$ from $p(\boldsymbol X \mid \boldsymbol\theta)$

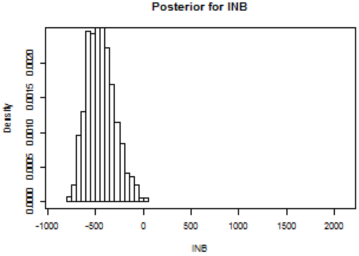


1. Select $q = 1, \ldots, Q$ values out of the S PSA simulations for $\boldsymbol{\theta}$

2. Simulate data $\boldsymbol{X}_q$ from $p(\boldsymbol{X} \mid \boldsymbol{\theta})$



3. Run the model to estimate $p(\boldsymbol{\theta} \mid \boldsymbol{X})$ and simulate values for $\text{INB}(\boldsymbol{\theta} \mid \boldsymbol{X}_q)$

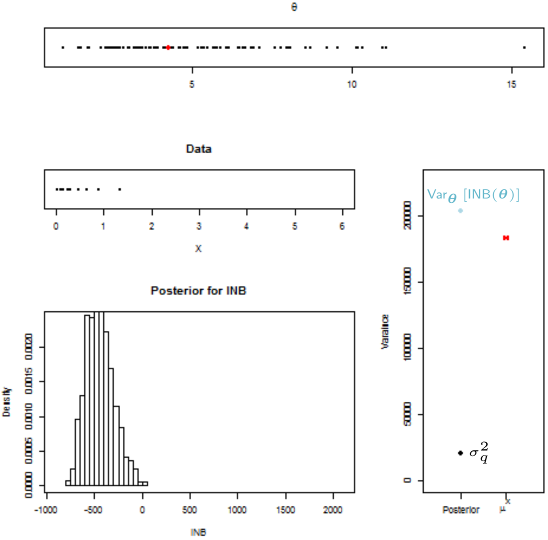


1. Select $q = 1, \ldots, Q$ values out of the S PSA simulations for $\boldsymbol{\theta}$

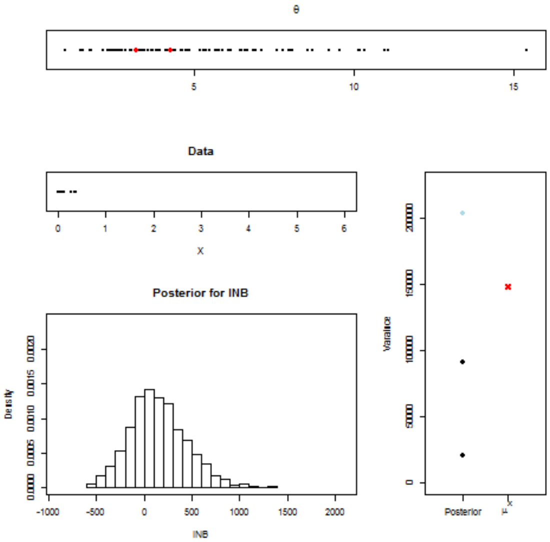
2. Simulate data $\boldsymbol{X}_q$ from $p(\boldsymbol{X} \mid \boldsymbol{\theta})$

3. Run the model to estimate $p(\boldsymbol{\theta} \mid \boldsymbol{X})$ and simulate values for $\text{INB}(\boldsymbol{\theta} \mid \boldsymbol{X}_q)$

4. Estimate the sample variance $\sigma_q^2 = \text{Var}_{\boldsymbol{\theta} \mid \boldsymbol{X}_q}[\text{INB}(\boldsymbol{\theta})]$



1. Select $q = 1, \ldots, Q$ values out of the S PSA simulations for $\boldsymbol\theta$
2. Simulate data $\boldsymbol X_q$ from $p(\boldsymbol X \mid \boldsymbol\theta)$
3. Run the model to estimate $p(\boldsymbol\theta \mid \boldsymbol X)$ and simulate values for $\text{INB}(\boldsymbol\theta \mid \boldsymbol X_q)$
4. Estimate the sample variance $\sigma^2_q = \text{Var}_{\boldsymbol\theta \mid \boldsymbol X_q}[\text{INB}(\boldsymbol\theta)]$

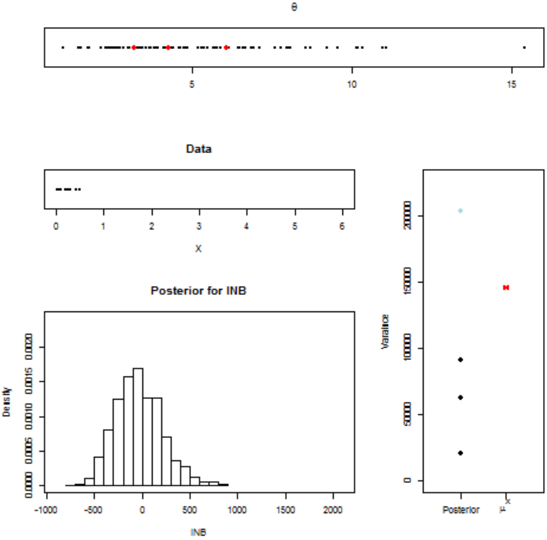


1. Select $q = 1, \ldots, Q$ values out of the S PSA simulations for $\boldsymbol{\theta}$

2. Simulate data $\boldsymbol{X}_q$ from $p(\boldsymbol{X} \mid \boldsymbol{\theta})$

3. Run the model to estimate $p(\boldsymbol{\theta} \mid \boldsymbol{X})$ and simulate values for $\text{INB}(\boldsymbol{\theta} \mid \boldsymbol{X}_q)$

4. Estimate the sample variance $\sigma^2_q = \text{Var}_{\boldsymbol{\theta} \mid \boldsymbol{X}_q}\left[\text{INB}(\boldsymbol{\theta})\right]$

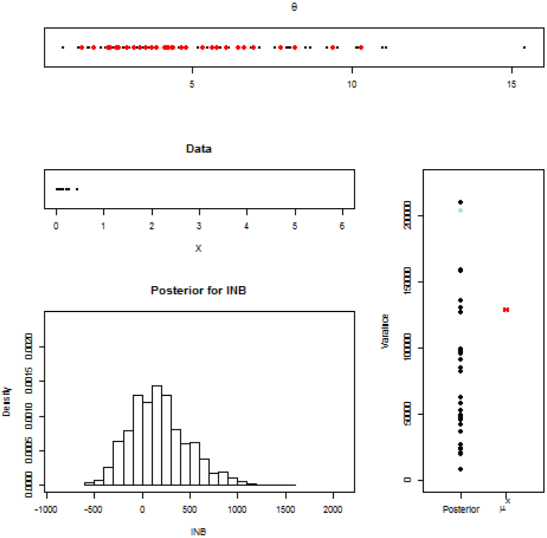


❶ Select $q = 1, \ldots, Q$ values out of the S PSA simulations for $\boldsymbol\theta$

❷ Simulate data $\boldsymbol X_q$ from $p(\boldsymbol X \mid \boldsymbol\theta)$

❸ Run the model to estimate $p(\boldsymbol\theta \mid \boldsymbol X)$ and simulate values for $\text{INB}(\boldsymbol\theta \mid \boldsymbol X_q)$

❹ Estimate the sample variance $\sigma^2_q = \text{Var}_{\boldsymbol\theta \mid \boldsymbol X_q}[\text{INB}(\boldsymbol\theta)]$



$$\underbrace{\text{Var}_{\boldsymbol X}\left[\text{E}_{\boldsymbol\theta \mid \boldsymbol X}[\text{INB}(\boldsymbol\theta)]\right]}_{\sigma^2_{\boldsymbol X}} = \underbrace{\text{Var}_{\boldsymbol\theta}[\text{INB}(\boldsymbol\theta)]}_{\sigma^2} - \underbrace{\text{E}_{\boldsymbol X}\left[\text{Var}_{\boldsymbol\theta \mid \boldsymbol X}[\text{INB}(\boldsymbol\theta)]\right]}_{\frac{1}{Q}\sum_{q=1}^{Q}\sigma^2_q}$$

❺ Use the Q estimates for σ^2_q to estimate the expected posterior variance

- Can now rescale the original PSA samples for INB$(\boldsymbol\theta)$ to ensure that mean and variance now match the computed values

$$\eta^{\boldsymbol X} = f(\mu^{\boldsymbol X}) = \text{INB}\left(\boldsymbol\theta^{(s)}\right)\sqrt{\frac{\sigma^2_{\boldsymbol X}}{\sigma^2}} + \mu\left(1-\sqrt{\frac{\sigma^2_{\boldsymbol X}}{\sigma^2}}\right)$$

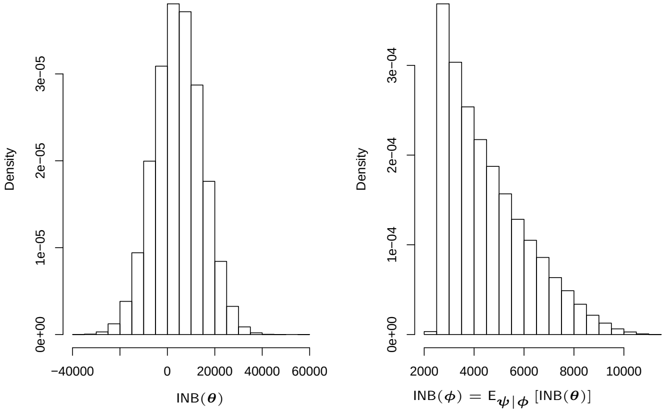
  - $\text{INB}\left(\boldsymbol\theta^{(s)}\right) = s-$th PSA simulation for the INB
  - $\mu = \text{E}_{\boldsymbol\theta}\left[\text{INB}(\boldsymbol\theta)\right] =$ PSA average INB
  - $\sigma^2 =$ PSA variance of the INB

and finally estimate the EVSI as

$$\text{EVSI} = \frac{1}{S}\sum_{s=1}^{S}\max\left\{0,\eta^{\boldsymbol X}\right\} - \max\{0,\mu\}$$

A Small Technicality...

- Only the focal parameters ϕ will be informed by the future study
- The distribution of $\mu^{\mathbf{X}}$ is similar to that induced by the EVPPI analysis!



- Can now rescale the original PSA samples for $\text{INB}(\boldsymbol\theta)$ to ensure that mean and variance now match the computed values

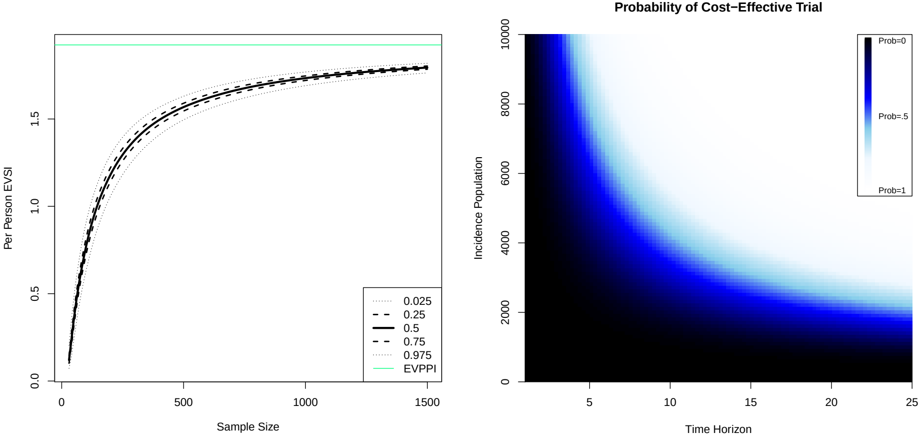
$$\eta^{\mathbf{X}} = f(\mu^{\mathbf{X}}) = \text{INB}\left(\boldsymbol\theta^{(s)}\right)\sqrt{\frac{\sigma^2_{\mathbf{X}}}{\sigma^2}} + \mu\left(1-\sqrt{\frac{\sigma^2_{\mathbf{X}}}{\sigma^2}}\right)$$

  - $\text{INB}\left(\boldsymbol\theta^{(s)}\right)$ = $s-$th PSA simulation for the INB
  - $\mu = \text{E}_{\boldsymbol\theta}\left[\text{INB}(\boldsymbol\theta)\right]$ = PSA average INB
  - σ^2 = PSA variance of the INB

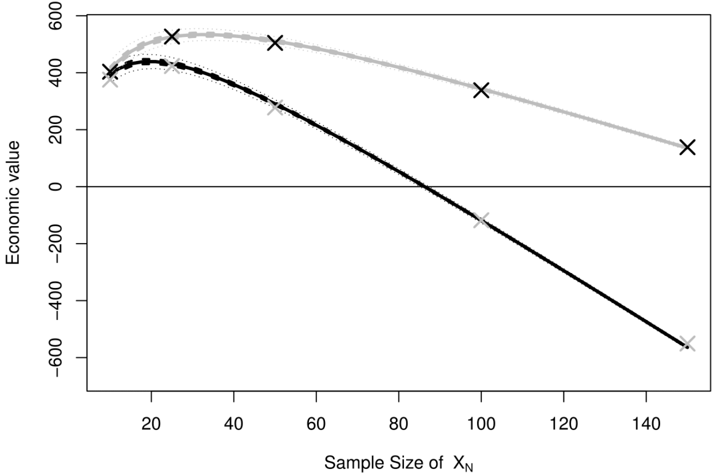
  and finally estimate the EVSI as

$$\text{EVSI} = \frac{1}{S}\sum_{s=1}^{S}\max\left\{0,\eta^{\mathbf{X}}\right\} - \max\{0,\mu\}$$

- Can also compute conditional version for $\boldsymbol\phi \in \boldsymbol\theta$. "Simply" substitute
  - σ^2 with $\sigma^2_{\boldsymbol\phi}$ = PSA variance for conditional INB (obtained using analysis of EVPPI)
  - $\text{INB}\left(\boldsymbol\theta^{(s)}\right)$ with $\text{INB}\left(\boldsymbol\phi^{(s)}\right) = \text{E}_{\boldsymbol\psi|\boldsymbol\phi}\left[\text{INB}\left(\boldsymbol\theta^{(s)}\right)\right]$



https://github.com/giabaio/EVSI
https://egon.stats.ucl.ac.uk/projects/EVSI
Heath et al *Value in Health*, 2018; **21(11)**: 1299-1304
Heath et al *Medical Decision Making*, 2017; **38(2)**: 163-173



Economic value
600
400
200
0
−200
−400
−600
20
40
60
80
100
120
140
Sample Size of X_N

Conclusions

- VoI can be very valuable in driving the whole economic evaluation process
  - Summarising PSA (in **addition** to standard tools, eg CEAC)
  - Research priority (in **place** of standard tools, eg sample size calculations?)

Conclusions

- VoI can be very valuable in driving the whole economic evaluation process
  - Summarising PSA (in **addition** to standard tools, eg CEAC)
  - Research priority (in **place** of standard tools, eg sample size calculations?)
- Historically limited use — also for computational complexity
  - Computation still a crucial component — but this is the price to pay for increasingly realistic and complex models?
  - Things can only get better(?) — recent research has improved this **massively**!

Conclusions

- VoI can be very valuable in driving the whole economic evaluation process
  - Summarising PSA (in **addition** to standard tools, eg CEAC)
  - Research priority (in **place** of standard tools, eg sample size calculations?)
- Historically limited use — also for computational complexity
  - Computation still a crucial component — but this is the price to pay for increasingly realistic and complex models?
  - Things can only get better(?) — recent research has improved this **massively**!
- Need standardised softward to enable practitioners to use the new tools
  - **And to move from Excel-based modelling to using fully proper statistical software (eg R):** http://www.statistica.it/gianluca/teaching/r-hta-workshop/2019/
  - Packages and web-applications exist to do this: `SAVI`, `BCEA`, `BCEAweb`, ...

Conclusions

- VoI can be very valuable in driving the whole economic evaluation process
  - Summarising PSA (in **addition** to standard tools, eg CEAC)
  - Research priority (in **place** of standard tools, eg sample size calculations?)
- Historically limited use — also for computational complexity
  - Computation still a crucial component — but this is the price to pay for increasingly realistic and complex models?
  - Things can only get better(?) — recent research has improved this **massively**!
- Need standardised softward to enable practitioners to use the new tools
  - **And to move from `Excel`-based modelling to using fully proper statistical software (eg `R`):** http://www.statistica.it/gianluca/teaching/r-hta-workshop/2019/
  - Packages and web-applications exist to do this: `SAVI`, `BCEA`, `BCEAweb`, ...
- Can we improve in any way by being more objective in our priors?
  - EVPPI with full GP based on conjugacy, while INLA/SPDE can get away with relatively flat priors (but has to fit LGM structure)
  - EVSI with Moment Matching has a bit more scope for subjective knowledge & informative priors — but (so far!) we've not experienced unwanted impact of priors...

Go raibh maith agat!

(**Thank you!**)