# Discussion on "Value of information" by Gianluca Baio (University College, London)

**Elías Moreno**

University of Granada, Spain
emoreno@ugr.es

OBayes, Warwick, June 28-July 2, 2019

The information contained in a sample is an elusive notion. The setting where the value of information is defined is the Cost–Effectiveness Analysis. The elements of CEA are:

- The net benefit of a treatment $t$,

$$z(e, c|R) = R \times e - c$$

  where $e$ is the random effectiveness of the treatment, $c$ is the random cost, and $R$ is a deterministic quantity whose meaning is the quantity the health provider is able to pay for the unit of effectiveness.
- The distribution $P_t(z|\theta, R)$, is the reward of treatment $t$, and $\theta$ an unknown parameter.
- The utility function of $z$ utilized by Gianluca is

$$U(z) = z,$$

  and hence the expectation $\mathbb{E}_{z|\theta,R}U(z) \equiv NB_t(\theta|R)$, is

$$NB_t(\theta|R) = R\,\mathbb{E}_{e|\theta}e - \mathbb{E}_{c|\theta}c.$$

## Expected Value of Partial Perfect Information (EVPPI).

1.- The parameter is split as $\theta = (\phi, \psi)$ and,

$$\pi(\theta) = \pi(\phi)\pi(\psi|\phi).$$

Then, the EVPPI is given by

$$\text{EVPPI} = \mathbb{E}_{\phi} \max_t \mathbb{E}_{\psi|\phi} NB_t(\theta|R) - \max_t \mathbb{E}_{\theta} NB_t(\theta|R) \geq 0.$$

- Is there a subparameter $\phi$ of interest in CEA?
- It is not the expectation of the net benefit the quantity of interest?

2.- Which the the role of $R$ in EVPPI?

3.- For a given $R$ and the set of competing treatments $\mathcal{T} = \{T_t, t = 1, \ldots, k\}$, the treatment given by

$$\max_{t=1,\ldots,k} \mathbb{E}_\theta NB_t(\theta, R)$$

is an optimal treatment with respect to the rewards

$$P_t(z|R) = \int P_t(z|\theta, R)\pi(\theta)\, d\theta,\ t = 1, \ldots, k.$$

- The treatment defined by

$$\mathbb{E}_\phi \max_t \mathbb{E}_{\psi|\phi} NB_t(\theta|R),$$

- is a treatment in $\mathcal{T}$ ?
- Which are the rewards?

4.- The parameters are always integrated out w.r.t. a prior. No data on $e$ and $c$ are used.

5.- I would like to have some properties of the EVPPI.

## Expected value of sampling information (EVSI)

1.- EVSI is defined for $x = (e, c)$ with distribution $P(x|\theta)$ as

$$\text{EVSI} = \left| \mathbb{E}_x \max_t \int NB_t(\theta|R)\pi(\theta|x)\ d\theta - \max_t \int NB_t(\theta|R)\pi(\theta)d\theta \right|$$

The treatment in $\{T_t, t = 1, \ldots, k\}$ from

$$\max_t \int NB_t(\theta|R)P_t(x|\theta)\pi(\theta)\ d\theta$$

depends on $x$. Then, is the treatment defined by

$$\int \left[ \max_t \int NB_t(\theta|R)P_t(x|\theta)\pi(\theta)\ d\theta \right]\ dx$$

in the class?
Remember that you assume that if the optimal treatment does not change the value of the information is cero.

## Expected value of sampling information (EVSI)

2.- Again, the definition of EVSI does not need data.

3.- The above expressions are computed from simulations, for instance, from the model

$$NB_t(\theta) = \mathbb{E}_{\psi|\phi} NB_t(\theta) + \varepsilon, \quad \varepsilon \sim N(\varepsilon|0, \sigma^2).$$

I have difficulties to understand this model.

**Thanks Gianluca for your work**

# Discussion on "Bayes–optimal frequentist intervals and tests" by Peter Hoff (Duke University)

**Elías Moreno**

University of Granada, Spain
emoreno@ugr.es

OBayes, Warwick, June 28-July 2, 2019

- Peter considers heterogenous independent normal sampling models

$$N(y_j|\theta_j, \sigma^2) \quad j = 1, \ldots, p,$$

where $\theta_1, \ldots, \theta_p$ are i.i.d. with prior distribution

$$N(\theta_j|\mu, \tau^2), \quad j = 1, \ldots, p.$$

- $\mathbf{y}_j = (y_{j1}, \ldots, y_{jn_j})$ is a sample of $y_j$ of size $n_j$.
- The interest is on estimating $\theta_j$.
- $\sigma$, $\mu$, $\tau$ are nuisance parameters for which no prior is assigned.

- The likelihood of $\sigma$, $\mu$, $\tau$ for the data $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_p)$, is

$$\ell(\sigma, \mu, \tau; \mathbf{y}) = \prod_{j=1}^{p} \int \left[ \prod_{i=1}^{n_j} N(y_{ji} | \theta_j, \sigma^2) \right] N(\theta_j | \mu, \tau^2) \, d\theta_j$$

  and $\hat{\sigma}^2$, $\hat{\mu}$ and $\hat{\tau}^2$ maximizes the likelihood. All data are here.
- An "empirical" Bayesian estimator of $\theta_j$ is the posterior expectation plus MLE,

$$\widehat{\theta}_j = \frac{n_j/\hat{\sigma}^2}{n_j/\hat{\sigma}^2 + 1/\hat{\tau}^2} \, \bar{y}_j + \frac{1/\hat{\tau}^2}{n_j/\hat{\sigma}^2 + 1/\hat{\tau}^2} \, \hat{\mu}.$$

- The MLEs do not bother me as they do a very good job for moderate or large sample sizes.

## Groups uncertainty

- The distribution $N(\theta_j|\mu, \tau^2)$ represents the uncertainty of $\theta_j$ around $\mu$ in group $j$, but it does not inform us about the uncertainty across groups.
- However, the smaller the $\hat{\tau}^2$ the bigger the influence of $\hat{\mu}$ on the estimator $\hat{\theta}_j$.
- Further, it is obvious that heterogeneity disappears if $\tau = 0$. Thus, for testing homogeneity we would testing the null hypothesis

$$H_0 : \tau = 0,$$

against

$$H_1 : \tau \geq 0.$$

- After testing we can proceed with either the homogeneous model (if $H_0$ is accepted) or the heterogeneous one (if $H_1$ is accepted)

- A difficulty arises when using $p-$values for testing:
  We cannot compute the probability of $H_0$ and $H_1$, and hence we do not know the weights to be given to the inference (estimation, etc.) from the null model

$$\Big\{N(y|\theta,\sigma^2),\pi(\theta)\Big\}$$

  and from the alternative model

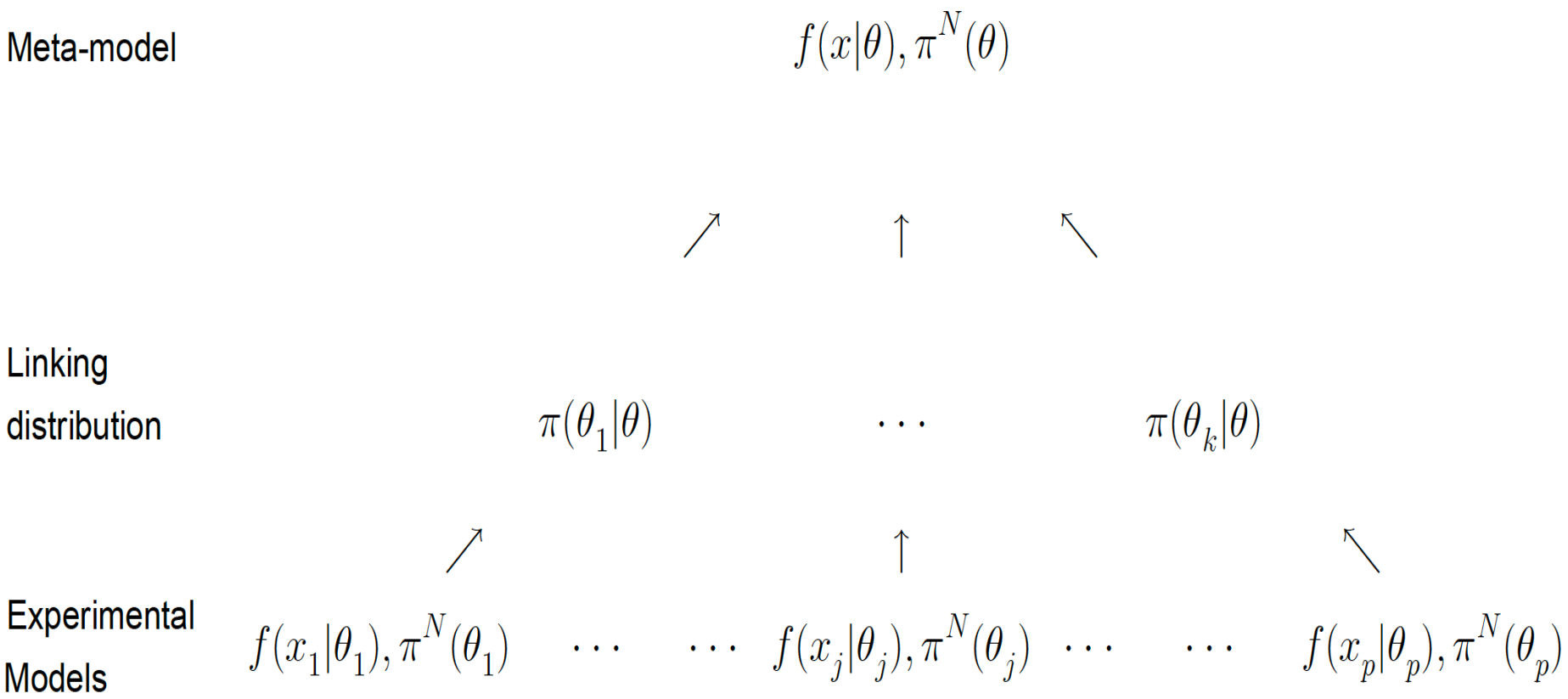
$$\Big\{N(y_j|\theta_j,\sigma^2),N(\theta_j|\mu,\tau^2)\quad j=1,\ldots,p\Big\}.$$

## Heterogeneity

Another related question is whether we should restrict ourself to fully heterogeneity (we note that partial heterogeneity is deemed possible)
Can we cluster some of the sampling models?

- Use clustering based on Bayesian model selection of the partition models (Hartigan 1990, Casella et al. 2014).
- The uncertainty on the cluster models must be incorporated into the final inference.

# Bayesian meta–analysis scheme




- $x$ is a latent variable (non observable)

## The prior distributions

- As the experimental model and the meta–model are the same, $\theta_j$ and $\theta$ are identically distributed with distribution $\pi^N(\cdot)$
- The linking distribution, $\pi(\theta_i|\theta)$ has to be compatible with $\pi^N(\cdot)$, that is

$$\int \pi(\theta_j|\theta)\pi^N(\theta)\, d\theta_j = \pi^N(\theta), \quad \int \pi(\theta_j|\theta)\pi(\theta)\, d\theta = \pi^N(\theta_j)$$

- $\pi(\theta_i|\theta)$ should be able to accommodate different degrees of concentration of $\theta_i$ around $\theta$

## A general linking distribution

From the model comparison between a generic experimental model

$$f(x|\theta_j), \pi^N(\theta_j)$$

and the meta–model for a fixed $\theta$

$$f(x|\theta)$$

the intrinsic prior $\pi^I(\theta_j|\theta, t)$ (Berger and Pericchi 1996, Moreno et al. 1998) satisfies:

i) $\pi^I(\theta_j|\theta, t)\pi^N(\theta)$ has marginals $\pi^N(\theta)$ and $\pi^N(\theta_j)$.

ii) Through $t$ we accommodate a given degree of concentration of $\theta_i$ around $\theta$

**Example**. Incidence data of wound infection in cesarean section for patients treated with prophylactic antibiotics versus placebo from seven randomized controlled trials (Sutton et al. 2002)

| Trial | Antibiotics | | Placebo | |
|---|---|---|---|---|
| | $n_i$ | $x_i$ | $n_i$ | $x_i$ |
| 1 | 100 | 3 | 33 | 0 |
| 2 | 11 | 1 | 17 | 5 |
| 2 | 42 | 0 | 40 | 0 |
| 4 | 167 | 4 | 140 | 5 |
| 5 | 36 | 1 | 25 | 1 |
| 6 | 115 | 12 | 117 | 15 |
| 7 | 16 | 0 | 16 | 1 |

Model: $Bin(x_i|\theta_i, n_i)$, $i = 1, \ldots, 7$. We are interested in estimating $\theta$, the meta–parameter.

# The top clusters

| Antibiotics | | Placebo | |
|---|---|---|---|
| Cluster Models | Post. Prob. | Cluster Models | Post. Prob. |
| {1, 3, 4, 5, 7}, {2}, {6} | 0.29 | {1}, {2}, {3}, {4}, {5}, {6}, {7} | 0.26 |
| {1}, {2}, {3}, {4}, {5}, {6}, {7} | 0.10 | {1, 3}, {2}, {4}, {5}, {6}, {7} | 0.04 |
| {1, 3, 4, 5}, {2}, {6}, {7} | 0.04 | {1, 3, 4, 5, 7}, {2, 6} | 0.0.3 |
| {1, 3, 4, 5, 6, 7}, {2} | 0.04 | the rest | $< 0.03$ |
| {1, 3, 4, 7}, {2}, {5}, {6} | 0.04 | | |
| the rest | $< 0.03$ | | |

# Posterior expectation of the meta–incidence of wound infection in cesarean

| Without clustering | | | After clustering | | |
|---|---|---|---|---|---|
| Antibiotics | Placebo | Incidence reduction | Antibiotics | Placebo | Incidence reduction |
| 0.024 | 0.045 | −0.021 | 0.057 | 0.070 | −0.013 |

Thanks Peter for you work