

A Bayesian semi-parametric approach for cost-effectiveness analysis in health economics

Un approccio bayesiano semiparametrico per l'analisi costi efficacia in economia sanitaria

Caterina Conigliani¹ and Andrea Tancredi²

¹ Dipartimento di Economia, Università di Roma Tre

² Dipartimento di Studi Geoeconomici, Università di Roma La Sapienza

e-mail: andrea.tancredi@uniroma1.it

Riassunto: Il confronto di due trattamenti tramite un'analisi costi efficacia risulta spesso problematico in quanto la distribuzione dei costi può presentare forte asimmetria, multimodalità ed elevata curtosi. Di conseguenza, poiché i modelli parametrici standard non riescono a cogliere contemporaneamente tutti questi aspetti, la stima dei costi medi legati ai due trattamenti risulta difficile. In questo lavoro condurremo un'analisi costi efficacia su un insieme reale di dati utilizzando un modello semiparametrico per i costi che consente di superare gli eventuali problemi di specificazione del modello.

Keywords: mean cost, mixture models, extreme value theory.

1. Introduction

When performing clinical trials it is increasingly common to evaluate the costs of the treatments, besides to their effectiveness, in order to perform a cost-effectiveness analysis. Indeed, such a kind of analysis, by quantifying the balance between effectiveness and cost of the treatments provides to the health service provider a useful tool to make policy decisions. Specifically, suppose that two treatments, T_1 and T_2 are to be compared in a randomised controlled trial, and that the available data are direct measurements of efficacy and cost

$$D = \{x_{ij} = (e_{ij}, c_{ij}) \quad i = 1, 2 \quad j = 1, 2, n_i\}$$

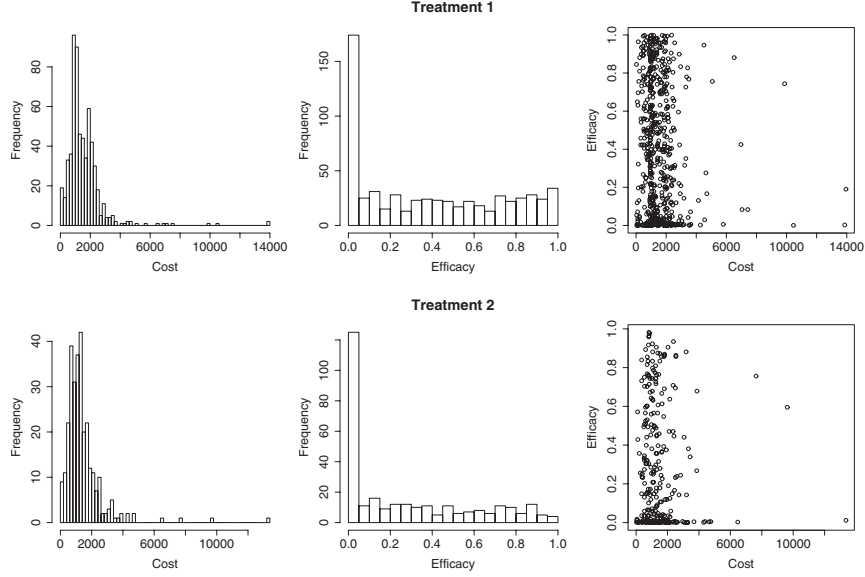
Let γ_i and μ_i be respectively the mean efficacy and the mean cost for the patients assigned to the treatments T_i . Moreover, let K be a decision makers's *willingness to pay* coefficient, that is the units of money a decision maker is prepared to pay to obtain one unit of effectiveness. To perform a cost-effectiveness analysis we could estimate the net benefit $K(\gamma_1 - \gamma_2) - (\mu_1 - \mu_2)$. In fact, if such a quantity is greater than 0, then T_1 is cost-effective relative to T_2 . A more general approach is obtained by evaluating the evidence for the hypothesis $K(\gamma_1 - \gamma_2) - (\mu_1 - \mu_2) > 0$ for different values of K . The resulting curve, called Cost Effectiveness Acceptability Curve (cfr. van Hout *et al* 1994), can be easily formalised in a Bayesian framework evaluating the posterior probability

$$Q(K) = P(K(\gamma_1 - \gamma_2) - (\mu_1 - \mu_2) > 0 | D)$$

for different values of K , as suggested by O'Hagan and Stevens (2001).

In this paper we present an example of cost-effectiveness analysis using real data. The data are the costs and the efficacies of two treatments. Figure 1 shows the data for each treatment.

Figure 1: Histograms and scatterplots for cost and efficacy of the treatments.



2. The model

Generally, cost data exhibit highly skew, heavy-tailed and multi-modal distributions. For this reason standard parametric models may not be adequate, as discussed, for example, in Nixon and Thompson (2004). Following Conigliani and Tancredi (2005) we adopt a Bayesian semi-parametric approach in which the bulk of the data and the tail are modelled separately. We model the observations below an unknown threshold with a mixture of unknown number of uniform distributions with unknown range. Specifically, the model for the bulk of cost data is

$$h(c|p^{(k)}, a^{(k)}, \alpha) = \sum_{l=1}^k p_l \mathcal{U}_{[a_l, a_{l+1})}(c).$$

Above the threshold we adopt the standard extreme value model for tail observations i.e. the GPD distribution

$$g(c|\alpha, \sigma, \xi) = \frac{1}{\sigma} \left(1 + \xi \frac{c - \alpha}{\sigma} \right)^{-1/\xi - 1}.$$

In this way, the resulting model for all the cost data

$$f(c) = \begin{cases} (1 - w) h(c|p^{(k)}, a^{(k)}, \alpha) & 0 < c < \alpha \\ w g(c|\alpha, \sigma, \xi) & \alpha \leq c < \infty \end{cases} \quad (1)$$

has the appealing property of catching all the possible relevant features of the data such as multimodality, asymmetry and kurtosis.

Calling $\theta_i = (k, p_{1,i}, \dots, p_{k,i}, a_{2,i}, \dots, a_{k,i}, \alpha_i, w_i, \sigma_i, \xi_i)$ the parameters of the Uniforms-GPD mixture model (1) for the cost distribution of the treatment T_i , it follows that the mean cost for such treatment is (if $\xi_i < 1$)

$$\mu_i = (1 - w_i) \sum_{l=1}^k p_{l,i} \frac{a_{l+1,i} + a_{l,i}}{2} + w_i \left[\alpha_i + \frac{\sigma_i}{1 - \xi_i} \right] \quad i = 1, 2$$

To model the effectiveness outcome is not problematic as to model the cost data. Since in our application the efficacy measure belongs to the interval $[0, 1]$ we take a Beta distribution. Moreover, to take account of possible correlation between cost and efficacy we suppose that, for treatment T_i

$$E(e_{ij}|c_{ij}) = \frac{\exp(\beta_{0i} + \beta_{1i}c_{ij})}{1 + \exp(\beta_{0i} + \beta_{1i}c_{ij})}$$

where β_{0i} and β_{1i} are unknown parameters.

It follows that the marginal mean efficacy for treatment T_i can be written as

$$\gamma_i = \int_0^\infty \frac{\exp(\beta_{0i} + \beta_{1i}c)}{1 + \exp(\beta_{0i} + \beta_{1i}c)} f(c|\theta_i) dc$$

and must be evaluated numerically.

Prior modelling for the parameters of the model for cost data is discussed in Conigliani and Tancredi (2005). Standard non informative priors are assumed for the parameters of the Beta distributions for the efficacy. Posterior inference is possible by MCMC methods.

3. Results

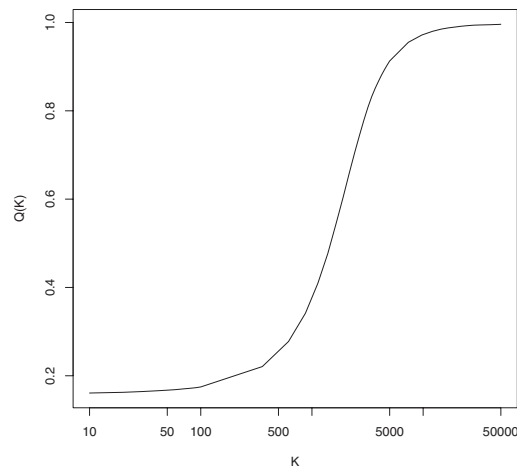
Table 1 shows posterior summaries for the difference between the mean costs and the mean efficacies. The posterior mean for the mean cost is higher for patients treated with T_1 and $P(\mu_1 - \mu_2 > 0|D) = 0.84$, indicating that the first treatment is more expensive. On the other hand, we have strong evidence that treatment T_1 is more effective. In fact, the credible interval for the difference $\gamma_1 - \gamma_2$ contains only positive values and $P(\gamma_1 - \gamma_2 > 0|D) = 0.99$.

Table 1: *Posterior summaries for the comparison of the two treatments*

	$\mu_1 - \mu_2$	$\gamma_1 - \gamma_2$
Posterior Mean	133	0.08
Credible Interval (0.95)	(-184,585)	(0.03,0.13)

Moreover, we computed the Cost-Effectiveness Acceptability Curve, that is plotted in Figure 2. We have that $Q(K) = 0.16$ at $K = 0$, corresponding to the probability that T_1 is cheaper, and tends to 0.99 as $K \rightarrow \infty$, corresponding to the probability that T_1 is more effective. Finally note that if the health provider were willing to pay about 5000 to obtain one unit of effectiveness, then there is a posterior probability of about 90 per cent that T_1 is more cost effective than T_1 . However for smaller values of K the case in favour of T_1 decreases rapidly.

Figure 2: *Cost-Effectiveness Acceptability Curve for the comparison of the two treatments.*



4. Discussion

We have considered the problem of assessing two treatments for their cost-effectiveness with a real example where data on both costs and efficacy are available from a clinical trial. Since conclusions from cost-effectiveness analysis are generally sensitive to the choice of the cost distribution, see Thompson and Nixon (2005), we have modelled the cost data with a semi-parametric model. However, other possible approaches to the problem, like sensitivity analysis with alternative distributions and Bayesian model averaging, can be helpful in drawing conclusions about cost-effectiveness.

References

- Conigliani C., Tancredi A. (2005) Semi-parametric modelling for costs of health care technologies, *Statistics in Medicine*, 24, 3171–3184.
- van Hout B. A., Al M. J., Gordon G. S., Rutten F. (1994) Costs, effects and C/E ratios alongside a clinical trial, *Health Economics*, 3 309–319
- O’Hagan A., Stevens J. W. (2001) A framework for cost-effectiveness analysis from clinical trial data, *Health Economics*, 10 303–315
- Nixon R.M., Thompson S. G. (2004) Parametric modelling of cost data in medical studies, *Statistics in Medicine* 23 1311–1331
- Thompson S. G., Nixon R.M. (2005) How sensitive are cost effectiveness analysis to the choice of parametric distributions? *Medical Decision Making* 25 416–423