European Journal of Operational Research 226 (2013) 173–182

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Interfaces with Other Disciplines

Optimal treatments in cost-effectiveness analysis in the presence of covariates: Improving patient subgroup definition

E. Moreno [a], F.J. Girón [b], M.L. Martínez [b], F.J. Vázquez-Polo [c,*], M.A. Negrín [c]

[a] Department of Statistics and O.R., University of Granada, 18071 Granada, Spain
[b] Department of Statistics and O.R., University of Málaga, 29071 Málaga, Spain
[c] Department of Quantitative Methods, U. Las Palmas de G.C., 35017 Las Palmas de G.C., Spain

ARTICLE INFO

Article history:
Received 8 May 2012
Accepted 5 November 2012
Available online 16 November 2012

Keywords:
Bayesian variable selection
Cost-effectiveness analysis
Bivariate regression model
Intrinsic priors
Subgroup analysis
Utility function

ABSTRACT

In the presence of covariates, the cost-effectiveness analysis of medical treatments shows that the optimal treatment varies across the patient population subgroups, and hence to accurately define the subgroups is a crucial step in the analysis. A patient subgroup definition using only influential covariates within the potential set of patients covariates established by the expert has recently been proposed, and the influential covariates were chosen from the univariate distributions of the effectiveness and the cost, conditional on the effectiveness. In this paper, we argue that the Bayesian variable selection procedure should be developed using the bivariate distribution of the cost and the effectiveness, which is not the usual practice. This new approach, provides results with wider applicability and more understandable without a significative increase in the complexity of the procedure. For real and simulated data sets, optimal treatments for subgroups are found, and compared with that from previous methods.

1. Introduction

Cost-effectiveness analysis refers to the statistical decision problem of choosing a medical treatment among m competitors by taking into account both the cost and the effectiveness of the treatments. This is an important optimization problem extensively treated in health economic literature (see, for instance, Barton et al. (2008) and references therein) although unfortunately underrepresented in operational research literature in spite of the papers by Brailsford and Harper (2008), Moreno et al. (2010b, 2012) and Pesch and Woeginger (2012).

The states of nature of this decision problem are the treatments net benefit random vector $\mathbf{z} = (z_1, \ldots, z_m)$, a m-dimensional vector in which the ith component, z_i, is defined as $z_i = R \cdot e_i - c_i$, where c_i and e_i are the cost and the effectiveness of treatment T_i, and R the utility assigned to a unit of effectiveness expressed in monetary units. We note that, conditional on a quantity R, which is fixed by the health provider, the distribution of the net benefit is determined by the bivariate distributions of the cost and the effectiveness of the treatments involved in the analysis.

Typically, the distribution of the cost and the effectiveness of the treatments $f_i(c_i, e_i|\theta_i, \mathbf{x})$, $i = 1, \ldots, m$, where $\theta_1, \ldots, \theta_m$ are unknown parameters and $\mathbf{x} = (x_1, \ldots, x_p)$ a set of p potential patient covariates, so the distribution of the net benefit $\mathbf{z}$ varies across patients. Therefore, the optimal treatment is a function of $\mathbf{x}$ and this motivates the subgroup analysis (Sculpher and Gafni, 2001; NICE, 2008; Espinoza et al., 2011). In addition, the optimal treatment dramatically varies across subgroups (Moreno et al., 2012), and hence to accurately define the existing patient population subgroups is of utmost importance. The underlying statistical problem of defining subgroups is the so-called variable selection problem.

Patients population subgroups have been defined in the health economic literature mostly using a statistical formulation that only permits the consideration of discrete covariates (Stinnett and Mullahy, 1998; Pocock et al., 2002; Willan et al., 2004; Nixon and Thompson, 2005; Sculpher, 2008; Manca et al., 2010; Gomes et al., 2012a,b, among others). In Moreno et al. (2012) the patient population subgroups were defined for covariates either continuous or discrete, and it was proposed using only a subset of influential covariates selected among the potential set of them with the help of a Bayesian variable selection procedure.

1.1. Motivation for a bivariate variable selection procedure

The bivariate sampling linear model for the cost and the effectiveness is typically represented as the product of a linear model for the effectiveness and a linear model for the cost conditional on the effectiveness, which now plays the role of a covariate, through the generic decomposition of the joint sample density

* Corresponding author. Tel.: +34 9 28 45 18 00; fax: +34 9 28458225.
E-mail addresses: emoreno@ugr.es (E. Moreno), fj_giron@uma.es (F.J. Girón), mlmartinez@uma.es (M.L. Martínez), fjvpolo@dmc.ulpgc.es (F.J. Vázquez-Polo), mnegrin@dmc.ulpgc.es (M.A. Negrín).

0377-2217/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2012.11.003

function $f(c,e|\theta,\mathbf{x}) = f(e|\theta_1,\mathbf{x})f(c|e,\theta_2,\mathbf{x})$, where θ, θ_1, and θ_2 are unknown parameters. Usually, the marginal sampling distribution for the effectiveness is assumed to be a normal regression, and for the cost, conditional on the effectiveness, was either a normal or a lognormal regression.

In Moreno et al. (2012), the variable selection was carried out using the decomposition of the bivariate cost-effectiveness sampling model mentioned above. By so doing univariate models $f(e|\theta_1,\mathbf{x}_{se})$ and $f(c|e,\theta_2,\mathbf{x}_{sc})$ were obtained, where $\mathbf{x}_{se}(\subset\mathbf{x})$ denotes the influential covariates selected for the effectiveness distribution and $\mathbf{x}_{sc}(\subset\mathbf{x})$ for the conditional distribution of the cost. On the other hand, let $f(c,e|\theta,\mathbf{x}_{sb})$ be the resulting model after applying the variable selection procedure to the bivariate distribution, where $\mathbf{x}_{sb}$ is the influential subset of covariates for the joint density.

The difficulty with these three models comes from the fact that the coherence equality

$$f(c,e|\theta,\mathbf{x}_{sb}) = f(e|\theta_1,\mathbf{x}_{se})\cdot f(c|e,\theta_2,\mathbf{x}_{sc})$$

does not necessarily hold, unless the subsets $\mathbf{x}_{sb}$, $\mathbf{x}_{se}$ and $\mathbf{x}_{sc}$ coincide.

Furthermore, even when the three subsets of covariates coincide it is not clear how the underlying uncertainty in choosing the covariates in the marginal density of the effectiveness is propagated to the conditional distribution of the cost, and hence it is not clear how to compute the total variable selection uncertainty. We note that the model space for the effectiveness, in which model selection is carried out, contains 2^p models while the model space for the cost, conditional on the effectiveness, contains 2^{p+1} models. Therefore, our uncertainties are set in different probability spaces making it difficult to evaluate the total uncertainty in the statistical variable selection procedure. This has undesirable implications. For instance, inferences based on model averaging cannot be considered when the covariates are selected from the univariate decomposition because of the weights are not well defined.

We remark that the subsets $\mathbf{x}_{sb}$, $\mathbf{x}_{se}$ and $\mathbf{x}_{sc}$ do not necessarily coincide as the following example shows. This example is based on a real clinical trial carried out in Hospital Clinic and Hospital de Bellvitge of Barcelona, Spain.

Example 1. This is an example based on a real data from a randomized clinical trial (Hernández et al., 2003) that compares two alternative treatments for exacerbated chronic obstructive pulmonary disease (COPD) patients. It was postulated that home hospitalization, treatment T_2, of selected chronic obstructive pulmonary disease exacerbations admitted at the emergency room could facilitate a better outcome than conventional hospitalization, treatment T_1. For patients under treatment T_2 integrated care was delivered by a specialized respiratory nurse with the patient's free-phone access to the nurse ensured for an 8-week follow-up period. We use information from 167 patients with COPD exacerbations over a 1-year period (1st November 1999 to 1st November 2000) among those admitted to the emergency department of two tertiary hospitals, Hospital Clinic and Hospital de Bellvitge of Barcelona, Spain. The two primary criteria for inclusion in the study were COPD exacerbation as a major cause of referral to the emergency room and absence of any criteria for imperative hospitalization as stated by the British Thoracic Society guidelines. The number of patients randomly allocated to treatment T_1 was 70 and 97 to T_2.

The six potential covariates considered in the clinical trial were the following: Age, Sex, Smoking habit, forced expiratory volume in one second (FEV), exacerbations requiring in-hospital admission (HOSV), and the score at the beginning of the study (SGRQ1).

Table 1 shows the variables selected for the bivariate normal model $f(c,e|\theta,\mathbf{x})$ and for the two-steps univariate procedure. The Bayesian variable selection procedure for the bivariate model is the one given in Section 2, and the Bayesian procedure for the univariate models is that given in Moreno et al. (2012).

Table 1
Influential variables for treatments T_1 and T_2, where e refers to the univariate linear regression model for effectiveness and $c|e$ denotes the linear regression model for cost given the effectiveness.

Model	Treatment T_1	Treatment T_2
Bivariate procedure	{SGRQ1, Age, FEV}	{SGRQ1, FEV}
Two-steps univariate conditioned procedure	e: {SGRQ1, Age}	e: {SGRQ1, FEV}
	$c|e$: {FEV}	$c|e$: {SGRQ1, Age}

The second row of Table 1 gives the influential variables for the treatments using the bivariate normal model, and in the third row the influential variables for the decomposed model. The conclusion we draw for treatment T_1 is that both ways of selecting variables convey the same message. However, for treatment T_2 the influential covariates for the bivariate model do not coincide with the ones selected when using the univariate models.

Remark 1. We are not saying that there is something wrong in the decomposition $f(c,e|\theta,\mathbf{x}) = f(e|\theta_1,\mathbf{x})\cdot f(c|e,\theta_2,\mathbf{x})$. Our concerns refer to the propagation of the uncertainty in the two-step variable selection procedure when using the decomposed model representation, and whether this coherence decomposition is fullfiled if $\mathbf{x}$ is changed in the right hand side of the equation.

To avoid the above difficulties we suggest returning to the original bivariate distribution $f(c,e|\theta,\mathbf{x})$ and selecting the covariates directly from this model. In this setting we have only one model space containing 2^p bivariate models and the above mentioned difficulties disappear.

The price we pay for defining subgroups with the variables selected for the bivariate variable (c,e) is the slightly higher complexity of the underlying sampling model that now becomes a 1×2 matrix-variate normal (or lognormal) distribution.

1.2. Notation

Let $f(\mathbf{y}|\theta,\mathbf{x})$ be the bivariate distribution of the vector of cost and effectiveness $\mathbf{y} = (c,e)^t$ of a generic treatment T, where θ is an unknown parameter vector, and $\mathbf{x} = (x_1,\ldots,x_p)^t$ a vector of p potential covariates. We denote by $f(z|R,\theta,\mathbf{x})$ the distribution of $z = R\cdot e - c$, the net benefit of the treatment, which is obtained from the distribution $f(\mathbf{y}|\theta,\mathbf{x})$. While the deterministic quantities R and $\mathbf{x}$ are observable, the parameters θ are unobservable and have to be eliminated in $f(z|R,\theta,\mathbf{x})$. We can do that with the help of a dataset $data = (\mathbf{y}_1,\ldots,\mathbf{y}_n,\mathbf{X})$, i.e. a sample of $\mathbf{y}$ of size n having a $n\times p$ design matrix $\mathbf{X}$.

In a frequentist setting, the unknown parameter $\theta$is replaced by the maximum likelihood estimator $\hat{\theta}(data)$ to obtain the predictive data dependent distribution $f(z|R,\hat{\theta}(data),\mathbf{x})$, and in the Bayesian approach the parameter $\theta$is assumed to be distributed following a prior distribution $\pi(\theta)$ that we update to the posterior distribution $\pi(\theta|data)$. Then, θ is eliminated by integration with respect to the posterior $\pi(\theta|data)$ to obtain

$$f(z|R,data,\mathbf{x}) = \int f(z|R,\theta,\mathbf{x})\pi(\theta|data)d\theta. \tag{1}$$

This predictive distribution of the net benefit (1) is the one to be used in the Bayesian cost-effectiveness decision problem. We note that under either the frequentist or the Bayesian viewpoint, the key distribution in the cost-effectiveness analysis is the posterior predictive density of the net benefit.