ROMA TRE
Università degli Studi di Roma Tre
Dipartimento di economia

Semi–parametric modelling
for costs of health care technologies

Caterina Conigliani* and Andrea Tancredi**

Working paper n. 34
2003

Aracne

00173 Roma, via Raffaele Garofalo, 133 A/B
tel. (06) 72672222 telefax (06) 72672233

www.aracne-editrice.it
info@aracne-editrice.it

ISBN 88-7999-574-X

I edizione: ottobre 2003

I Working Papers del Dipartimento di Economia svolgono la funzione di divulgare tempestivamente, in forma definitiva o provvisoria, i risultati di ricerche scientifiche originali. La loro pubblicazione è soggetta all'approvazione del Comitato Scientifico.
Per ciascuna pubblicazione vengono soddisfatti gli obblighi previsti dall'art.1 del D.L.L. 31.8.1945, n. 660 e successive modifiche.
Copie della presente pubblicazione possono essere richieste alla Redazione

Università degli Studi di Roma Tre – Dipartimento di Economia
Via Ostiense, 139 – 00154 Roma
Tel. 06-57374003 Fax 06-57374093 E-mail: dip_eco@uniroma3.it

* Università Roma Tre, Italy.
** Università di Roma "La Sapienza", Italy.

Index

Abstract. *Cost data that arise in the evaluation of health care technologies usually exhibit highly skew, heavy-tailed and, possibly, multi-modal distributions. Distribution-free methods for analysing these data, such as the bootstrap, or those based on the asymptotic normality of sample means, may often lead to inefficient or misleading inferences. On the other hand, parametric models that fit the data (or a transformation of the data) equally well can produce very different answers. We consider a Bayesian approach, and model cost data with a distribution composed of a piecewise constant density up to an unknown endpoint, and a generalised Pareto distribution for the remaining tail.*

1 Introduction

In cost evaluations conceived to have an impact on medical policy, the main interest is the total healthcare cost, so that it is inference on population mean costs that is informative. However, cost data obtained for individual patients in health economic studies typically exhibit highly skew and heavy tailed distributions, and many problems arise with the various approaches currently available for analysing such data.

In fact, as discussed in O'Hagan and Stevens (2002, 2003), nonparametric methods, such as those based on the asymptotic normality of the sample mean or nonparametric bootstrapping, may be inefficient and their justification breaks down in small samples. On the other hand, parametric modelling may lead to more efficient inference, but is dependent on the population distribution matching the model adequately. The main difficulty in this sense is that the high skewness and kurtosis usually found in cost data imply that the population mean can be very sensitive to the tail of the distribution beyond the range of the data; one consequence of this is that parametric models that fit the data equally well can produce very different answers. Another problem related to the parametric modelling of cost data concerns possible transformations of the data; in fact, as discussed in Thompson and Barber (2000) and Briggs and Gray (1998), mean values and confidence limits may be difficult to interpret on the transformed scales, and back-transformation onto the original scale is not always straightforward.

Here we consider a Bayesian approach, and for the simple problem of estimating the population mean cost from data on a single sample of patients, we model the bulk of the data and the tails separately. More specifically, we consider a distribution composed of a piecewise constant density up to an unknown endpoint, and a generalised Pareto distribution (GPD) for the remaining tail. Note this model has been applied to environmental data by Tancredi *et al.* (2002).

The first component of the model, the step function, is very flexible, in the sense that it has the appealing property of catching all the relevant features of the data; if for instance the data exhibit multimodality, the corresponding model will be multimodal. However, the step function will hardly give any weight to values beyond the range of the data; for this reason, we introduce a different model for the right tail of the distribution, the GPD, that is often used in extreme value theory to model tail data (see, for instance Coles, 2001). Recall that extreme value theory is mainly concerned with quantifying the stochastic behaviour of a process at unusually large (or small) levels, and provides a class of models to enable estimation of the probability of events that are more extreme than any that have already been observed.

The model and notations are introduced in Section 2. In Section 3 we apply our model on a real data set originating from a study comparing two treatments for asthma, and estimate the population mean cost under both treatments. A few concluding remarks are presented in the final section.

2 The model

Let $h(x\,|\cdot)$ be a piecewise constant density on $(0,\alpha)$ with unknown number of steps s at positions $a_1 = 0 < a_2 < \ldots < a_s < a_{s+1} = \alpha$ taking value ω_i on the subinterval $[a_i, a_{i+1})$; if $a^{(s)} = (a_2, \ldots, a_s)$ denotes the vector of unknown step positions and $\omega^{(s)} = (\omega_1, \ldots, \omega_s)$ denotes the vector of unknown heights, we can write:

$$h\left(x\,\middle|\,s, \omega^{(s)}, a^{(s)}, \alpha\right) = \sum_{i=1}^{s} \omega_i I_{[a_i, a_{i+1})}(x) \tag{1}$$

with the constraint

$$\sum_{i=1}^{s} \omega_i \left(a_{i+1} - a_i\right) = 1.$$

The step function (1) has been analysed in a Bayesian context by Robert (1998) and Robert and Casella (1999), and can be seen as a mixture of s uniform distributions $U_{[a_i, a_{i+1})}$:

$$h\left(x\,\middle|\,s, p^{(s)}, a^{(s)}, \alpha\right) = \sum_{i=1}^{s} p_i U_{[a_i, a_{i+1})} \tag{2}$$

where $p^{(s)} = (p_1, ..., p_s)$, $p_i = \omega_i\,(a_{i+1} - a_i)$, and $\sum_{i=1}^{s} p_i = 1$.

Moreover, let $g\,(x\,|\alpha, \sigma, \xi\,)$ be the generalized Pareto density with threshold α, scale parameter σ and shape parameter ξ, all unknown:

$$g\,(x\,|\alpha, \sigma, \xi\,) = \frac{1}{\sigma}\left[1 + \frac{\xi\,(x-\alpha)}{\sigma}\right]^{-\frac{1}{\xi}-1}, \tag{3}$$

defined on

$$\left\{x : x > \alpha \text{ and } 1 + \frac{\xi\,(x-\alpha)}{\sigma} > 0\right\}.$$

Then we model cost data $x_1, ..., x_n$ as *i.i.d.* observations from a distribution with density

$$f\big(x\big|s, p^{(s)}, a^{(s)}, \alpha, \sigma, \xi, \omega\big) = \begin{cases} (1-\omega)\,h\big(x\,\big|s, p^{(s)}, a^{(s)}, \alpha\big) & 0 < x < \alpha \\ \omega\,g(x\,|\alpha, \sigma, \xi) & \alpha \leq x < \infty \end{cases} \tag{4}$$

where ω is the probability that an observation is greater than α.

Note that the qualitative behaviour of the GPD depends significantly on the value of the the shape parameter ξ. In particular, $g\,(x\,|\alpha, \sigma, \xi)$ is an increasing function of x when $\xi < -1$, is constant when $\xi = -1$, while is a decreasing function of x for $\xi > -1$; moreover, it has an upper bound of $\alpha - \sigma/\xi$ if $\xi < 0$, while it has no upper limit if $\xi \geq 0$. For this reason, when modelling cost data such as those arising in the evaluation of health care technologies it seems appropriate to impose the constraint $\xi \geq 0$. Furthermore, the GPD has finite expected value

$$E\,(x\,|\alpha, \sigma, \xi\,) = \alpha + \frac{\sigma}{1-\xi}$$

only if ξ is less than 1; as our main interest is the population mean μ, in what follows we will also impose the constraint $\xi < 1$, so that

$$\mu = (1-\omega)\sum_{i=1}^{s} p_i \frac{a_{i+1} + a_i}{2} + \omega\left[\alpha + \frac{\sigma}{1-\xi}\right]. \tag{5}$$

The behaviour of the GPD for $\xi \in [0, 1)$ is shown in Figure 1; $g\,(x_i\,|\alpha, \sigma, \xi)$ decreases exponentially for $\xi = 0$ and polinomially for $\xi > 0$, and the tail gets heavier as ξ increases.

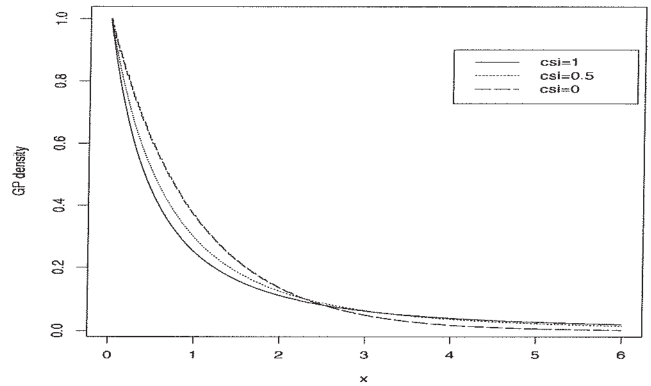

Figure1: Generalized Pareto density for different values of ξ ($\alpha = 0$, $\sigma = 1$)

Moreover, it is important to notice that when approximating the distribution of a tail with the GPD, the scale parameter σ depends on the threshold α (see Coles, 2001 §4.1), and one should take this into account when introducing a prior distribution for the parameters of the model. However, as pointed out in Tancredi *et al.* (2002) and in Coles and Tawn (1996), this problem can easily be avoided by reparametrizing model (4) in terms of the parameters of the generalized extreme value distribution (GEV) corresponding to the GPD (Coles, 2001). In particular we can write σ and ω as:

$$\sigma = \psi + \xi\,(\alpha - m)$$

$$\omega = \frac{1}{n}\left[1 + \frac{\xi\,(\alpha - m)}{\psi}\right]^{-\frac{1}{\xi}}$$

where m and ψ are the location and scale parameters of the GEV, ξ is the shape parameter of the GEV, that coincides with the shape parameter of the GPD, α is the threshold and n is the sample size; this parametrization has the advantage of making the parameters m, ψ, ξ independent of α.

Finally, note that Bayesian inference for this model is possible using *Markov chain Monte Carlo* (MCMC) methods (see, for instance, Robert and Casella, 1999), and details can be found in Tancredi *et al.* (2002).

3 Example: the IMPACT study

We present an example using a study comparing Pulmicort Turbuhaler® and Azmacort® in the treatment of asthma. The design of the study was a randomised, multicentre, open-label parallel group study; a total of 945 patients (out of which 905 were included in the health economic evaluation) from 26 Managed Care Organizations were randomised in a 2:1 ratio to the Pulmicort Turbuhaler® and Azmacort® treatment groups respectively, and were followed for one year starting from their date of randomisation. The primary clinical endpoint was symptom-free days, while economic endpoints were both annual asthma-specific healthcare costs and annual total costs (asthma-specific plus asthma-associated healthcare costs).

Note that, as the problem we are considering is that of estimating a mean cost, here we will study only cost data; Table 1 shows sample summaries for asthma-specific costs and for total costs under both treatments.

Table 1: IMPACT study: sample descriptive statistics

	Pulmicort Turbuhaler®		**Azmacort®**	
	asthma-specific costs	total costs	asthma-specific costs	total costs
sample size	608	608	297	297
mean	1060.70	1297.94	985.92	1168.07
st. dev.	864.58	1242.19	866.53	968.44
median	869.12	1039.51	815.74	939.25
minimum	17.83	17.83	13.79	13.79
maximum	9553.61	15349.42	9151.25	9253.41
skewness	4.88	5.30	4.61	3.58
kurtosis	40.62	44.29	36.10	23.61

Consider first asthma-specific cost data. As shown in Table 1, patients treated with Pulmicort Turbuhaler® have a higher mean cost compared to patients treated with Azmacort®. For the rest, the two distributions seem to have similar features: for both treatment groups the standard deviations are large, indicating that the data are spread quite far around the mean, and the median cost is smaller than the mean, indicating positively skew data; this fact is confirmed also by the standard skewness statistic $\overline{\mu}_3/\sigma^3$. Finally, in both cases the kurtosis statistic $\overline{\mu}_4/\sigma^4$ indicate that the two distributions of asthma-specific costs are significantly leptokurtic.

Consider now total cost data. Again, as shown in Table 1, patients treated with Pulmicort Turbuhaler® have a higher mean cost compared to patients treated with Azmacort®, and for both treatment groups the distribution of costs is highly skew and leptokurtic. However, for patients treated with Pulmicort Turbuhaler® the standard deviation is much larger compared to patients treated with Azmacort®, and the maximum cost is 65% larger than the maximum Azmacort® cost. In fact, two patients treated with Pulmicort Turbuhaler® had particularly large total costs, namely \$10075.49 and \$15349.42, so that in this case the distribution of costs is more skew and more leptokurtic than when we considered only asthma-specific costs, as emphasised also by the skewness and the kurtosis statistics (note that, on the contrary, for patients treated with Azmacort® the skewness and kurtosis statistics decrease when we move from asthma-specific costs to total costs).

3.1 Estimating asthma-specific mean costs

We now apply our model to the IMPACT data in order to estimate and compare the asthma-specific mean costs under the two treatments. First, we need complete the models by introducing the prior distributions for the unknown parameters. In particular, we assume that for both treatment groups the parameters of the step function are independent of the remaining parameters of the model. Moreover, we express weak prior informations by introducing a uniform prior on the range of the data for the threshold α, a uniform prior between 1 and a maximum number of steps $s_{\max} = 20$ for s, and a Dirichlet distribution with all parameters set equal to 1, *i.e.* a uniform prior over the region $\sum_{i=1}^{s} p_i = 1$, for the weights $p^{(s)}$. Then, following Green (2000), as a prior distribution for the step positions $a^{(s)}$ we consider a joint density providing a slight preference against two step positions occurring too closely in succession:

$$\pi\left(a^{(s)} \,|s, \alpha\right) \propto a_2\left(a_3 - a_2\right) \dots \left(\alpha - a_s\right). \tag{6}$$

Finally, we explore the behaviour of three different prior distributions for the GEV parameters. First, in a totally non-informative case, *a priori* we model the GEV parameters in the two treatment groups independently, and introduce the standard non-informative prior

$$\pi\left(m, \psi, \xi\right) = \pi\left(m\right) \pi\left(\psi\right) \pi\left(\xi\right) \propto \frac{1}{\psi}. \tag{7}$$

It is interesting to notice that in this case the posterior mean of the mean cost μ is infinite; this can be seen by looking at $g(x_i|\alpha, \sigma, \xi)$ and μ as functions of ξ, and is due to the fact that at $\xi = 1$ $g(x_i|\alpha, \sigma, \xi)$ is bounded and non-zero:

$$g(x_i|\alpha, \sigma, \xi) \stackrel{\xi\to 1}{\longrightarrow} \left(1 + \frac{x_i - \alpha}{\sigma}\right)^{-2}$$

while μ is unbounded, so that

$$\int_0^1 \mu \prod_i g(x_i|\alpha, \sigma, \xi)\, \pi(\xi)\, d\xi = \infty.$$

Clearly this problem could have been avoided by introducing a different prior for ξ, for instance a beta distribution with parameters (c, d) such that $c \geq 1$ and $d \geq 2$; for this reason, we explore also the behaviour of a $Beta(2, 2)$, so that

$$\pi(m, \psi, \xi) = \pi(m)\, \pi(\psi)\, \pi(\xi) \propto \frac{\xi(1-\xi)}{\psi}. \tag{8}$$

Finally, note that although we might have little prior information, we would not expect the tails of the distributions of asthma-specific costs to be very different between the two treatments. In fact, as argued in O'Hagan and Stevens (2003), in general the large costs observed under one treatment are indicative of the extreme skewness and kurtosis of costs generally, and their presence suggest that in larger samples one might find comparable skewness and kurtosis also under other treatments. Accordingly, we consider also an informative Bayesian analysis, in which the GEV parameters under the two treatments are modelled as exchangeable by introducing the hierarchical priors:

$$\begin{array}{lll} m \sim N(\varepsilon_1, \eta_1) & \log(\psi) \sim N(\varepsilon_2, \eta_2) & \log\left(\frac{\xi}{1-\xi}\right) \sim N(\varepsilon_3, \eta_3) \\ \varepsilon_1 \sim uniform & \varepsilon_2 \sim uniform & \varepsilon_3 \sim N(0, 1) \end{array} \tag{9}$$

where ε_1, ε_2 and ε_3 are estimated using data from both treatment groups, and η_1, η_2 and η_3 are known constants. In particular, as we require that the prior distributions on m, ψ and ξ individually are non-informative, but also that all three parameters *a priori* have similar values in the two treatment groups, here we have assumed $\eta_1 = 1000^2$, $\eta_2 = 0.4^2$ and $\eta_3 = 1$. In fact, under these values the ratio of the scale parameters tends to be not too far from unity, and the differences of the location and of the shape parameters tend to be not too far from zero (accounting for the parameters' scales). Note that m is the location parameter of the GEV, *i.e.* of the sample maximum distribution, so that in general $\sqrt{\eta_1}$ should be of the same order of magnitude of the expected difference of maximum costs, when this information is available *a priori*, or at least of the observed difference of maxima. Finally, note that in this case is

$$\pi(\xi) = \int \pi(\xi \mid \varepsilon_3)\, \pi(\varepsilon_3)\, d\varepsilon_3 \propto \frac{1}{\xi(1-\xi)} \exp\left\{ -\frac{1}{2(1+\eta_3)} \left[\log\left(\frac{\xi}{1-\xi} \right) \right]^2 \right\}$$

and it is straightforward to prove that the posterior mean of the mean cost μ is finite.

For each of the three prior distributions discussed for the GEV parameters, and for both treatment groups, Table 2a, Table 2b and Table 2c show a 95% posterior credible interval ($PCI_{0.95}$) and a point estimate for some of the parameters of the model (the posterior median for μ and the posterior mean for α, ξ, s), obtained with 100000 iterations of the simulation algorithm discussed in Tancredi *et al.* (2002).

Table 2a: Asthma-specific costs: posterior summaries in the totally non-informative case

	Pulmicort Turbuhaler®		Azmacort®	
	point estimate	$\mathbf{PCI_{0.95}}$	**point estimate**	$\mathbf{PCI_{0.95}}$
μ	1088	$(1008, 1408)$	1003	$(913, 1197)$
α	1282	$(1248, 1524)$	940	$(708, 1243)$
ξ	0.60	$(0.38, 0.86)$	0.37	$(0.14, 0.67)$
s	6.05	$(5, 9)$	2.27	$(2, 4)$

Table 2b: Asthma-specific costs: posterior summaries assuming a Beta (2,2) prior for ξ

	Pulmicort Turbuhaler®		Azmacort®	
	point estimate	$\mathbf{PCI_{0.95}}$	**point estimate**	$\mathbf{PCI_{0.95}}$
μ	1081	$(1005, 1280)$	1006	$(913, 1192)$
α	1268	$(1248, 1309)$	947	$(707, 1263)$
ξ	0.58	$(0.38, 0.81)$	0.38	$(0.17, 0.67)$
s	5.98	$(5, 8)$	2.31	$(2, 4)$

Table 2c: Asthma-specific costs: posterior summaries assuming an exchangeable prior for the GEV parameters

	Pulmicort Turbuhaler®		Azmacort®	
	point estimate	$PCI_{0.95}$	point estimate	$PCI_{0.95}$
μ	1051	$(995, 1120)$	1029	$(934, 1166)$
α	1281	$(1247, 1685)$	934	$(734, 1207)$
ξ	0.50	$(0.33, 0.66)$	0.43	$(0.23, 0.63)$
s	6.00	$(5, 8)$	2.27	$(2, 4)$

Consider first the results in Table 2a, obtained assuming the prior distribution (7). As expected, the posterior mean of the shape parameter ξ is higher for patients treated with Pulmicort Turbuhaler® than under Azmacort®, indicating that the tail of the distribution of asthma-specific costs is heavier for the first treatment group. Moreover, the posterior median of the mean cost μ is higher for patients treated with Pulmicort Turbuhaler® than under Azmacort®, indicating that the first treatment is more expensive (recall that in this case the posterior mean of μ is infinite). We also looked at the posterior distribution of the difference of the mean costs for patients treated with Turbuhaler® and for patients treated with Azmacort®, that we denote with $\Delta_\mu = \mu_1 - \mu_2$; the posterior median, a 95% posterior credible interval, and the posterior probability that Δ_μ is greater than zero were

$$Me\,(\Delta_\mu \,|\, \underline{x}) = 88, \quad PCI_{0.95} = (-119;\ 415)\,, \quad P\,(\Delta_\mu > 0 \,|\, \underline{x}) = 0.85$$

respectively.

Consider now the results in Table 2b, obtained assuming the prior distribution (8); recall that in this case the posterior mean of μ is finite, and for the two treatment groups we obtained 1097 and 1018 respectively. Comparing Table 2b with Table 2a, we notice that here the posterior distributions of ξ under the two treatments have been pulled towards each other. In fact, for patients treated with Pulmicort Turbuhaler® the posterior mean of ξ has decreased, and so has the upper limit of the credible interval for ξ; at the same time, for patients treated with Azmacort® the posterior mean of ξ has increased, and so has the lower limit of the credible interval for ξ. This is clearly the effect of the $Beta\,(2, 2)$ prior, that gives more prior weight to values around 0.5. As a consequence, also the posterior distributions of the mean cost under the two treatments have been pulled towards each other. In fact, looking at the distribution of Δ_μ, we found

$$Me(\Delta_\mu|\underline{x}) = 78, \;\; E(\Delta_\mu|\underline{x}) = 79, \;\; PCI_{0.95} = (-124; 289), \;\; P(\Delta_\mu > 0 \,|\, \underline{x}) = 0.83.$$

Consider now the results in Table 2c, obtained assuming the prior distribution (9). Comparing Table 2b with Table 2c we notice that the exchangeable prior for the GEV parameters has more effect than the $Beta\,(2,2)$ prior in terms of shrinkage; this is due to the fact that the former suggests that all the GEV parameters, not just ξ, are expected *a priori* to have similar values in the two treatment groups. Here the posterior mean of the mean cost μ for the two treatment groups were 1053 and 1034 respectively, and the posterior summaries of Δ_μ were

$$Me(\Delta_\mu|\underline{x}) = 21, \;\; E(\Delta_\mu|\underline{x}) = 19, \;\; PCI_{0.95} = (-88; 115), \;\; P(\Delta_\mu > 0 \,|\, \underline{x}) = 0.65.$$

It is also interesting to look at the posterior density estimates under the two treatment groups, that were found to have very similar behaviour assuming either of the prior distributions considered for the GEV parameters. Figure 2 and Figure 3 show the posterior density estimates, together with the histograms of the data, for the totally non-informative case.

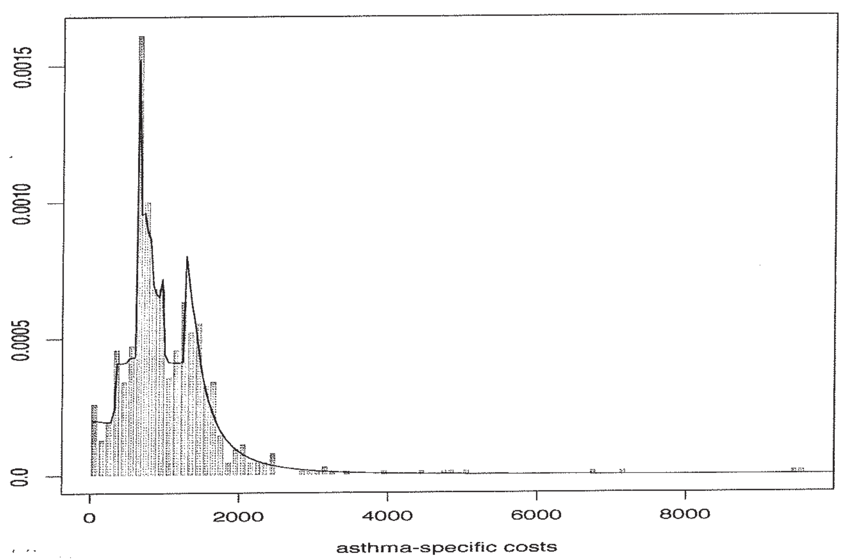

Figure 2: Pulmicort Turbuhaler®: histogram and posterior density estimate

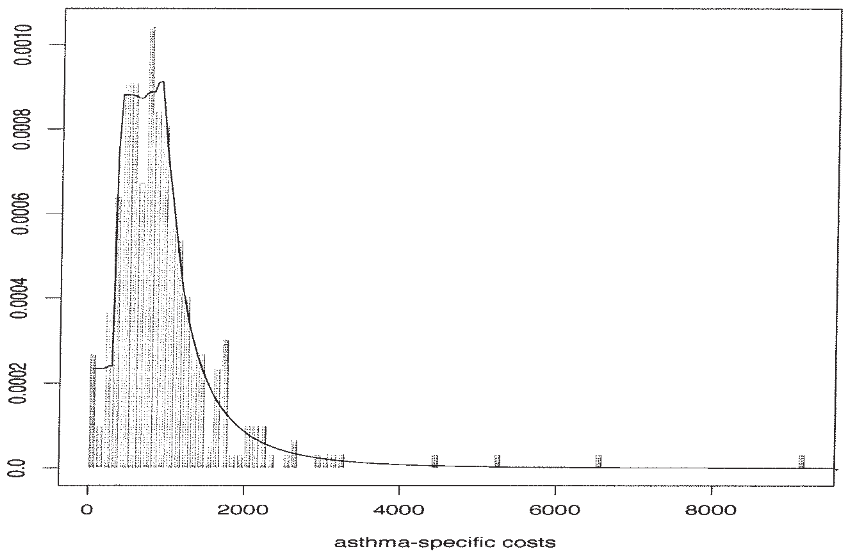

Figure 3: Azmacort®: histogram and posterior density estimate

These estimates were obtained by computing at each iteration the density (4) on a grid of points, and by averaging over the number of iterations. In both cases the fitting of the model is quite good. In particular, we notice that the asthma-specific costs for patients treated with Pulmicort Turbuhaler® exhibit bimodality, and the posterior density estimate has the same behaviour; note that this particular feature of these data explains why the posterior expected number of steps was quite large in this case, compared with the posterior expected number of steps for patients treated with Azmacort®.

3.2 Comparisons with other methods

As pointed out in the Introduction, cost data usually exhibit highly skew and heavy-tailed distributions, and it can be extremely difficult to produce realistic probabilistic models for the underlying population distribution. For this reason, non-parametric statistical methods, that may be applied without regard to the shape of the population distribution, are particularly attractive. These methods are quite common in the frequentist setting, and include the approach based upon assuming that the mean of a (large) sample of costs will approximately follow a normal distribution (see, for instance, Thompson and Barber, 2000) and the non-parametric bootstrap (see, for instance, Briggs and Gray, 1998); note that both methods are based on sample means, and are justified only in large samples: under the standard normal-theory a confidence interval for a population mean is based on the Student t distribution, while the non-parametric bootstrap consists of drawing a large number of independent random samples with replacement from the original data, computing the mean of each sample and considering the bootstrap distribution that is produced. A Bayesian analogue of the non-parametric bootstrap was suggested by Rubin (1981), but in the Bayesian setting there seem to be more emphasis on the formulation of appropriate parametric models for cost data. For instance, in recognising the extreme skewness and kurtosis of costs, Al and van Hout (2000) assumed an underlying lognormal distribution, and O'Hagan and Stevens (2001) used a lognormal distribution for performing a cost-effectiveness analysis.

Here, for purposes of comparison, we tackle the problem of estimating asthma-specific costs using both the non-parametric approaches discussed above and a Bayesian parametric approach with a lognormal distribution. For the Bayesian model, following O'Hagan and Stevens (2001), we consider both a non-informative analysis and an informative analysis in which, in order to moderate the influence of the extreme sample values, we express a prior belief that the variances of log-costs should not be too different between treatment groups. The results are presented in Table 3.

Table 3: Asthma-specific costs: point estimates and interval estimates obtained with different methods

	Pulmicort Turbuhaler®		**Azmacort®**	
	Estimate	**95% Interval**	**Estimate**	**95% Interval**
Normal theory	1061	(992; 1129)	986	(887; 1084)
Non-parametric bootstrap	1060	(996; 1130)	987	(893; 1092)
Lognormal model (weak prior)	1104*	(1036; 1181)	1024*	(930; 1135)
Lognormal model (informative prior)	1104*	(1035; 1181)	1022*	(930; 1132)

* posterior median.

Three points emerge clearly from Table 3. First, the normal analysis and the non-parametric bootstrap give very similar results, which is often the case even when the skewness is very strong, as pointed out for instance in O'Hagan and Stevens (2003). Second, the informative prior in the lognormal model does not seem to have much of an effect here; this is probably due to the fact that on the log scale the variances of asthma-specific costs under the two treatments are quite similar (0.54 and 0.58 respectively), so that the prior distribution is not adding much information. Third, both the non-parametric frequentist approaches and the Bayesian parametric analysis produce smaller interval estimates than those presented in Table 2a, Table 2b and Table 2c, with the frequentist intervals being concentrated at smaller values of the mean cost. This is an interesting point: looking at the data in Figure 2 and Figure 3, it is obvious that asthma-specific costs are non-normal; the fact that more realistic assumptions like the lognormal distribution or our semiparametric model lead to completely different inferences seem then to contradict the claim that the normal theory and bootstrap methods are appropriate no matter what the underlying population distribution. On the other hand, whilst it is obvious that the lognormal distribution is a more appropriate assumption than the normal distribution, in general the true distributions of costs are unlikely to be lognormal or any other standard distributional form. In this sense our semi-parametric approach seems particularly attractive. In the IMPACT example, for instance, the fact that the upper limit of the posterior credible interval obtained with our model is always larger than the one obtained with the lognormal distribution seems to show that the latter tend to underestimate the weight of the right tail of the distribution of asthma-specific costs.

3.3 Estimating total mean costs

Recall that economic endpoints in the IMPACT trial were both annual asthma-specific healthcare costs and annual total costs. The analysis of total costs, however, produced results in accordance with those shown in Section 3.1, so that here we do not give much details of it. One point that is worth pointing out concerns the distribution of costs for patients treated with Pulmicort Turbuhaler®; recall that two of them had very large total costs, so that the skewness and the kurtosis statistics were even larger compared to the asthma-specific costs case. However, assuming either of the prior distributions (7), (8) and (9), we noticed that when moving from asthma-specific costs to total costs, the posterior mean of the shape parameter ξ did not change, neither did the posterior credible interval for ξ. This is an interesting point: it is showing that those two large observations are totally coherent with our model, they are not outliers, so that they do not increase the estimate of the tail shape. Indeed, what the two large observations modify is the estimate of the mean costs μ under the two treatments: in fact, looking at the distribution of Δ_μ, we found

$$Me\,(\Delta_\mu \,|\, \underline{x}) = 154, \quad PCI_{0.95} = (-52;\ 549)\,, \quad P\,(\Delta_\mu > 0 \,|\, \underline{x}) = 0.94$$

under the uniform prior,

$$Me\,(\Delta_\mu|\underline{x}) = 158,\ E\,(\Delta_\mu|\underline{x}) = 166,\ PCI_{0.95} = (-55; 467)\,,\ P\,(\Delta_\mu > 0 \,|\, \underline{x}) = 0.93$$

under the $Beta\,(2,2)$ prior, and

$$Me\,(\Delta_\mu|\underline{x}) = 62,\ E\,(\Delta_\mu|\underline{x}) = 57,\ PCI_{0.95} = (-128; 213)\,,\ P\,(\Delta_\mu > 0 \,|\, \underline{x}) = 0.77$$

under the exchangeable prior (with $\eta_1 = 10000^2$, $\eta_2 = 0.4^2$ and $\eta_3 = 1$).

4 Conclusions

We have presented a semi-parametric approach for modelling the cost data that arise in the evaluation of health care technologies in order to make inference on population mean costs. The model combines the semi-parametric approach to density estimation based on mixture models and the semi-parametric approach to tail estimation based on extreme value theory. The result is a very flexible model able to fit data set with very different shapes both in the bulk of data and in the tail. In particular, this model is suited for cost data which, in fact, exhibit higly skew, heavy-tailed and, possibly, multi-modal distributions.

One drawback of our model is related to the well known problems that may arise in conducting Bayesian inference using a continuous sampling distribution in presence of repeated observations due to rounding (see, for instance, Fernandez and Steel, 1999). In fact, consider n *i.i.d.* replications $\underline{x} = (x_1, x_2, ..., x_n)$ from model (4), and let $a_1 = 0 < a_2 < ... < a_s < a_{s+1} = \alpha$ be the step positions that identify the piecewise constant density (1). Then the posterior density can be written as

$$\pi\left(a^{(s)}, p^{(s)}, s, \alpha, m, \psi, \xi|\underline{x}\right) = \pi\left(a^{(s)}|\, p^{(s)}, s, \alpha, m, \psi, \xi, \underline{x}\right)\, \pi\left(p^{(s)}, s, \alpha, m, \psi, \xi|\underline{x}\right)$$

where, from (2) and (6), is

$$\pi\left(a^{(s)}|\, p^{(s)}, s, \alpha, m, \psi, \xi, \underline{x}\right) \propto \prod_{i=1}^{s}\left(\frac{p_i}{a_{i+1}-a_i}\right)^{n_i}(a_{i+1}-a_i),$$

and n_i is the number of observations falling between a_i and a_{i+1} (with $\sum_{i=1}^{s} n_i < n$). It follows that if two step positions a_j and a_{j+1} tend to assume the same value x, then $\pi\left(a^{(s)}|\, p^{(s)}, s, \alpha, m, \psi, \xi, \underline{x}\right)$ goes to infinity whenever $n_j > 1$, *i.e.* if more than one observation in the sample are recorded as having the same value x. Clearly, this problem could be avoided by introducing a different prior for the step positions, *i.e.* a distribution of the form

$$\pi\left(a^{(s)}\,|s, \alpha\right) \propto (a_2)^{f-1}(a_3-a_2)^{f-1}\ldots(\alpha-a_s)^{f-1} \qquad (10)$$

with f greater than the maximum number of repeated observations in the sample. However, in (10) the hyperparameter f controls the prior beliefs on the length of the steps, in the sense that by increasing f we favour steps of equal length. It follows that large values of f result in a much less flexible model. For this reason, in the analysis of IMPACT data we avoided the problem of repeated observations by slightly perturbing the original data, adding a uniformly distributed random number on (-1,1) to the recorded costs (note that the maximum number of repeated observations for patients treated with Pulmicort Turbuhaler® and for patients treated with Azmacort® were 20 and 5 respectively in the case of asthma-specific costs, and 9 and 5 respectively for total costs).

In conclusion, we believe that the novelty of our approach is that we can report model-based inference for mean costs without having to be too concerned about model misspecification problems. Moreover, by embedding our model in a two dimensional distribution, our approach can be used also to model jointly costs and efficacy measures, in order to carry out a cost-effectiveness analysis.

Acknowledgements

We are grateful to John Stevens and to AstraZeneca R&D Charnwood, UK, for providing data to use as examples, and to Tony O'Hagan for some usefull comments and ideas. We would also like to acknowledge helpful discussions at the CHEBS Focus Fortnight on "Costs", Sheffield (UK), January 27 - February 7, 2003.

References

Al, M.J. and van Hout, B.A. (2000) A Bayesian approach to economic analyses of clinical trials: the case of stenting versus balloon angioplasty. *Health Economics*, 9, 599-609.

Briggs, A. and Gray, A. (1998) The distribution of health care costs and their statistical analysis for economic evaluation. *J. Health Serv. Res. Policy*, 3, 233-245.

Coles, S.G. (2001) *An Introduction to Statistical Modeling of Extreme Values*, Springer-Verlag, London.

Coles, S.G. and Tawn, J.A. (1996) A Bayesian analysis of extreme rainfall data. *Appl. Stat.*, 45, 463-478.

Fernandez, C. and Steel, M.F.J. (1999). On the Dangers of Modelling Through Continuous Distributions: a Bayesian Perspective. In *Bayesian Statistics 6*, edited by J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith, Oxford University Press.

Green, P.J. (2000) A Primer on Markov Chain Monte Carlo. In *Complex Stochastic Systems*, edited by O.E. Barndorff-Nielsen, D.R. Cox, C. Kluppelberg, Chapman & Hall.

O'Hagan, A. and Stevens, J.W. (2001) A framework for cost-effectiveness analysis from clinical trial data. *Health Economics*, 10, 302-315.

O'Hagan, A. and Stevens, J.W. (2002) Bayesian methods for design and analysis of cost-effectiveness trials in the evaluation of health care technologies. *Statistical Methods in Medical Research*, 11, 469-490.

O'Hagan, A. and Stevens, J.W. (2003) Assessing and comparing costs: how robust are the bootstrap and methods based on asymptotic normality? *Health Economics*, 12, 33-49.

Robert, C.P. (1998) *Discretization and MCMC convergence assessment.* Lecture Notes in Statistics 135, Springer-Verlag, New York.

Robert, C.P. and Casella, G. (1999) *Monte Carlo Statistical Methods* Springer-Verlag, New York.

Rubin, D.B. (1981) The Bayesian bootstrap. *Ann. Stat.*, 9, 130-134.

Tancredi, A. Anderson, C. and O'Hagan, A. (2002) A Bayesian model for threshold uncertainty in extreme value theory. *Working paper 2002.12, Dipartimento di Scienze Statistiche, Università di Padova*, Submitted.

Thompson, S.G. and Barber, J.A. (2000) How should cost data in pragmatic randomised trials be analysed? *Br. Med. J.*, 329, 1197-1200.

WORKING PAPER
PUBBLICATI

1997

1. M. D'Antonio, M. Scarlato
   Struttura economica e commercio estero: un'analisi per le province italiane
2. P. Garegnani, A. Palumbo
   Accumulation of capital
3. E. Cerrito
   Crisi di cambio e problemi di politica monetaria nell'Italia di fine Ottocento. Appunti su alcune evidenze empiriche

1998

4. F. Manni
   Struttura delle fonti di finanziamento: un'indagine sulle principali società italiane produttrici di bevande
5. J.O. Berger, J. Mortera
   Default Bayes factors for one-sided hypothesis testing
6. A. Trezzini
   Capacity utilisation in the log run: a reply to Serrano
7. A.M. Ferragina
   Quality product differentiation in CEE-EUIntra-Industry trade
8. M.F. Renzi, L. Cappelli, G. Salerno
   Outsourcing: opportunità e limti per le aziende che operano con sistemi di qualità conformi alle norme ISO 9000
9. M. Scarlato
   The impact of international trade on employment and wage differentials: some evidence from the italian macro-regions
10. A. Trezzini
    Some notes on long-run capacity utilisation, steady state and induced investment

1999

11. S. Monni
    A convergence analysis of human development
12. G. M. Rey
    Informazione e politiche pubbliche: non è mai troppo tardi

2000

13. M.M. Barbieri, C. Conigliani
Fractional bayes factors for the analysis of autoregressive models with possible unit roots
14. M. Scarlato
Capitale sociale e Sviluppo Economico
15. A.M. Ferragina
Price versus quality competition in Italy's trade with Central and Eastern Europe over Transition
16. M. D'Antonio, M. Scarlato
Capitale umano e sviluppo economico. Un modello di equilibrio economico generale per il Centro–Nord ed il Mezzogiorno d'Italia
17. M. Cenci, L. Foffo Ciucci
Gli effetti della regolamentazione sull'attività di Insider Trading
18. A. Gheno
Alberi binomiali e struttura della volatilità
19. M. Cenci
Modelli per la struttura a termine con volatilità stocastica (una rassegna critica)
20. M. Cenci, A. Gheno
Metodologie per la valutazione delle obbligazioni convertibili in ipotesi di evoluzione stocastica della struttura per scadenza

2001

21. M. D'Antonio
L'interdipendenza tra economia ed ambiente: un'analisi applicata all'Italia
22. A. Naccarato
Uno studio sulla diffusione di fenomeni metereologici mediante particolari modelli spazio–temporali
23. M. Cenci, M. Scarlato
Settore sommerso e politiche di emersione: un approccio stocastico
24. S. Fadda
Per un nuovo sostegno allo sviluppo locale: patti territoriali e servizi al territorio nel nuovo scenario economico

2002

25. F. De Filippis, D. Storti
Le politiche di sviluppo rurale nell'Unione Europea: un "secondo pilastro" tutto da inventare

26.F. De Filippis, L. Salvatici
The eastwards Enlargement of the European Union and the Common Agricultural Policy

27.S. Fiumara, E. Bellisario
La rilevanza della certificazione di qualità nella responsabilità del produttore di giocattoli

28.S. Fadda
Does the change of Economic institutions require a change in values?

29.A. Corese
La mediateca: approcci gestionali

30.S.M. Fratini
Il teorema dell'indice e l'unicità dell'equilibrio con produzione: una discussione dei limiti di procedimento

31.C. Conigliani, F. Spezzaferri
An alternative Bayes factor for testing for unit autoregressive roots

32.M. Ciurlia
Applicazione di tecniche di simulazione alla valutazione delle opzioni russe

2003

33.C. D'Apice
La distribuzione del reddito in Italia nell' ultimo trentennio

34.C. Conigliani, A. Tancredi
Semi–parametric modelling for costs of healt care technologies

Finito di stampare nel mese di ottobre del 2003
dalla tipografia « Grafica Editrice Romana» di Roma
per conto della « Aracne editrice S.r.l. » di Roma