Chapman & Hall/CRC Biostatistics Series

Bayesian Methods in Health Economics

Gianluca Baio

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

A CHAPMAN & HALL BOOK

1

Introduction to health economic evaluation

This chapter is written by Rachael M. Hunter (Department of Primary Care and Population Health, UCL) and Gianluca Baio

1.1 Introduction

In recent years health economics has become an increasingly important discipline in medical research, especially with the transition from the paradigm of *evidence based medicine* to that of *translational research* (Berwick, 2005; Lean et al., 2008), which aims at making basic research applicable in the context of real practice, and under budget constraints, in order to enhance patients' access to optimal health care.

Since the 1970s, health care services have undergone dramatic changes: increasing demand for health care has generated an increase in the number of available interventions, which have sometimes been applied regardless on considerations about the actual quality and the costs associated.

Consequently, decision-makers responsible for the provision of health care are increasingly facing critical appraisal processes of the modality in which they manage the available resources and they need to adjust the management and the evaluation of the processes used, with respect to some measures of clinical benefit. The main reasons for the necessity of containing cost associated with health care are essentially the following:

- The progressive increase of the proportion of the "older" (above 65 years) population;
- the increase of life expectancy and of the incidence of chronic and degenerative pathologies;
- the refinement of diagnostic techniques;
- the availability of innovative health technologies and therapeutic tools associated with better clinical outcomes but also with higher costs.

In this perspective, the systematic analysis of organised data provides a fundamental contribution to the identification of economically appropriate

strategies. This in turns has helped the integration between several clinical and quantitative (*e.g.* statistics and economics) disciplines, so much that it can be reasonably argued that *health economics* is in fact a combination of medical research, epidemiology, statistics and economics. Figure 1.1 shows this concept graphically and highlights the fact that health economics encompasses more than the mere cost evaluations.

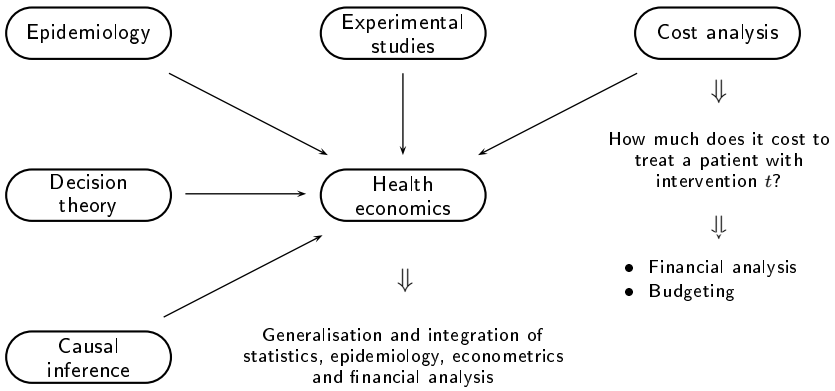


FIGURE 1.1
Health economic evaluation as the integration of different disciplines. Cost analysis only represents one side of the story

In this chapter we present some concepts that are relevant to the definition and description of health economic evaluations. In particular we discuss the characteristics of the two dimensions along which economic evaluations are conducted: costs and clinical benefits. The latter can be defined in several different ways, each of which gives rise to a specific method of analysis. We present the main ones in §1.6. Finally, we give a first introduction to the problem of comparative evaluation of two or more health interventions, which will be discussed in more technical detail in chapters 3 and 5.

1.2 Health economic evaluation

Health economics can be formally described as the application of economic theory to *health* (defined as "a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity"; WHO, 2012) and *health care*, *i.e.* the diagnosis, treatment and prevention of disease and

TABLE 1.4
An example of calculation of QALYs from utility scores. In the two treatment groups ($t = 0, 1$), the measurements consist of the utility score u_{tj} and the costs c_{tj}, for $j = 0, \ldots, 4$ occasions

	Baseline	**6 months**	**12 months**	**18 months**	**24 months**	**Total**
Treatment group ($t = 1$)						
Utility score	0.656	0.744	0.85	0.744	0.744	
QALYs		0.350	0.399	0.399	0.372	1.519
Costs		£2 300	£300	£300	£300	£3 200
Control group ($t = 0$)						
Utility score	0.656	0.656	0.656	0.656	0.744	
QALYs		0.328	0.328	0.328	0.350	1.334
Costs		£300	£300	£300	£300	£1 200
Difference						
in QALYs ($\mathrm{E}[\Delta_e]$)[a]						0.185
in costs ($\mathrm{E}[\Delta_c]$)[a]						£2 000
Cost per QALY[a]						£10 811

[a] These quantities are defined in §1.7

where

$$\delta_j = \frac{\text{time between measurements } j \text{ and } (j-1), \text{ in years}}{1 \text{ year}}$$

For example, the QALYs at 6 months (*i.e.* at time $j = 1$) for treatment $t = 0$ are computed as

$$q_{01} = \left(\frac{0.656 + 0.656}{2}\right)\left(\frac{0.5}{1}\right) = 0.164,$$

since the time between the two measurements, 6 months, is only half a year. Similarly, for $t = 1$ the computation gives

$$q_{11} = \left(\frac{0.656 + 0.744}{2}\right)\left(\frac{0.5}{1}\right) = 0.35.$$

Overall, for each treatment the QALYs can be computed by summing the q_{tj} terms across all the time periods. In the present example, the measurements are repeated at every 6 month intervals and thus all values are added up over the 2 years. We define the QALYs using the notation e_t (to indicate the "effectiveness" of the treatment) as

$$e_t = \sum_{j=1}^{J} q_{tj}.$$

This produces a result of 1.519 extra QALYs for the patient under $t = 1$ and only 1.334 extra QALYs for the single patient under $t = 0$.

Notice that, in more realistic cases, instead of a single patient per group, we would have access to a sample of patients and therefore the relevant measures would be the population average computed across all relevant individuals. □

Each of the three main types of economic evaluation described above has strengths and weaknesses, and although each has their own specific characteristics most economic evaluations generally combine aspects of each. More in depth information can be found in Drummond et al. (2005). NICE's Decision Support Unit also has extensive guidance to support technical appraisals `http://www.nicedsu.org.uk/`. Table 1.5 summarises the main differences among them.

TABLE 1.5
A comparison of the characteristics of the main types of economic evaluation. Adapted from Meltzer and Teutsch (1998)

Type	**Costs included**[a]		**Outcomes**
	Direct	*Indirect*	
Cost-benefit	✓	✓	Monetary unit
Cost-effectiveness	✓	often	Health outcome[b]
Cost-utility	✓	rarely	Utility measure[c]

[a] All future costs and benefits should be discounted to the reference year (cfr. §1.5)
[b] For example: number of deaths averted
[c] For example: QALYs (cfr. §1.6.4)

1.7 Comparing health interventions

As discussed earlier, the purpose of economic evaluations is to provide information to decision-makers about the costs and outcomes of health care options to help with resource allocation decisions. Generally, economic summaries are computed in the form of "cost-per-outcome" ratios.

Moreover, in order to compare the two interventions ($t = 0, 1$), we can define suitable incremental population summaries, such as the population average *increment in benefits*, suitably measured as utilities (as in a CUA) or by means of hard clinical outcomes (as in a CEA):

$$\mathrm{E}[\Delta_e] = \overline{e}_1 - \overline{e}_0 \tag{1.2}$$

2

Introduction to Bayesian inference

2.1 Introduction

In the context of statistical problems, the frequentist (or empirical) interpretation of probability has played a predominant role throughout the twentieth century, especially in the medical field. In this approach, probability is defined as the limiting frequency of occurrence in an infinitely repeated experiment.

The underlying assumption is that of a "fixed" concept of probability, which is unknown but can be theoretically disclosed by means of repeated trials, under the same experimental conditions. Moreover, probability is generally regarded in classical statistics as a physical property of the object of the analysis.

However, although the frequentist approach still plays the role of the standard in various applied areas, there are many other possible conceptualisations of probability characterising different philosophies behind the problem of statistical inference. Among these, an increasingly popular is the Bayesian (sometimes referred to as subjectivist, in its contemporary form), originated by the posthumous work of Reverend Thomas Bayes (1763) and the independent contributions by Pierre Simone Laplace (1774, 1812) — see Howie (2002), Senn (2003), Fienberg (2006) or Bertsch McGrayne (2011) for a historical account of Bayesian statistics.

The main feature of this approach is that probability is interpreted as a subjective degree of belief in the occurrence of an event, representing the individual level of uncertainty in its actual realisation (cfr. de Finetti, 1974, probably the most comprehensive account of subjective probability). One of the main implications of subjectivism is that there is no requirement that one should be able to specify, or even conceive of some relevant sequence of repetitions of the event in question, as happens in the frequentist framework, with the advantage that "one-off" type of events can be assessed consistently (Dawid, 2005).

In the Bayesian philosophy, the probability assigned to any event depends on the individual whose uncertainty is being expressed and on the state of background information in light of which this assessment is being made. As any of these factors changes, so too might the probability. Consequently, under the subjectivist view, there is no assumption of a unique, correct (or "true")

value for the probability of any uncertain event. Rather, each individual is entitled to their own subjective probability and according to the evidence that becomes sequentially available, they tend to update their belief.

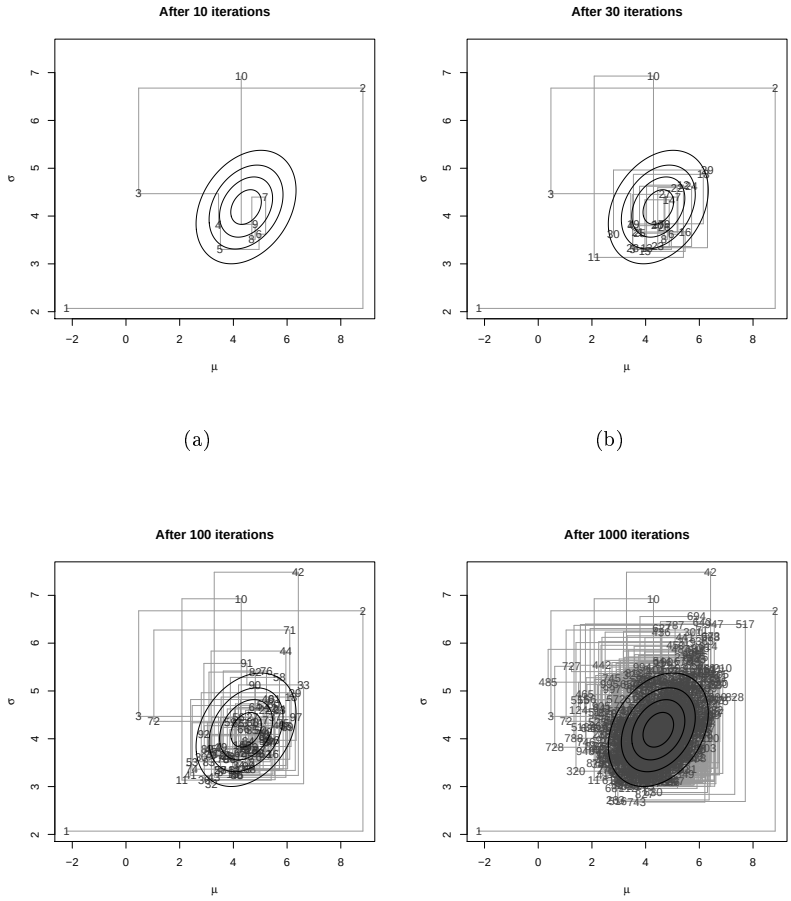
The development of Bayesian applied research has been limited probably because of the common perception among practitioners that Bayesian methods are "more complex". In fact, in our opinion the apparent higher degree of complexity is more than compensated by at least the two following consequences. First, Bayesian methods allow taking into account, through a formal and consistent model, all the available information, *e.g.* the results of previous studies. Moreover, the inferential process is straightforward, as it is possible to make probabilistic statements directly on the quantities of interest (*i.e.* some unobservable feature of the process under study, typically — but not necessarily — represented by a set of parameters).

In our opinion, Bayesian methods allow the practitioner to make the most of the evidence: in just the situation of "repeated trials", after observing the outcomes (*e.g.* successes and failures) of many past trials (and no other collateral information), all subjectivists will be drawn to an assessment of the probability of obtaining a success on the next event that is extremely close to the observed proportion of successes so far. However, if past data are not sufficiently extensive, it may reasonably be argued that there should indeed be scope for interpersonal disagreement as to the implications of the evidence. Therefore the Bayesian approach provides a more general framework for the problem of statistical inference.

Justifications for the use of Bayesian methods in health care evaluation have been detailed by Spiegelhalter et al. (2004), in terms of the formal quantitative inclusion of external evidence in all aspects of clinical research, including design, analysis, interpretation and policy-making. In particular, the Bayesian approach is valuable because: *i)* it proves more flexible and capable of adapting to each unique situation; *ii)* it represents a more efficient inferential tool, making use of all available evidence and not restricting formal evaluations to just the current data at hand; *iii)* it is particularly effective in producing predictions and inputs for decision-making.

Jackman (2009) suggests that performing statistical analysis by means of Bayesian methods also produces advantages from a pragmatic point of view: because of the wide availability of fast and cheap computing power, simulation-based procedures have allowed researchers to exploit more and more complex statistical models, especially under the Bayesian paradigm. Examples include the possibility of computing interval estimations in a straightforward way, without the need to rely on asymptotic arguments. This in turns has the potential of rendering "hard statistical problems easier".

The account of Bayesian statistics that is presented in this chapter is far from exhaustive — more comprehensive references are O'Hagan (1994), Berry (1996), Bernardo and Smith (1999), Lindley (2000), Robert (2001), Lee (2004), Spiegelhalter et al. (2004), Gelman et al. (2004), Lindley (2006), Lancaster (2008), Carlin and Louis (2009), Jackman (2009) and Christensen et al. (2011).



(a) (b)

(c) (d)

FIGURE 2.10
Gibbs sampling simulation for the semi-conjugated Normal model. The numbers indicate the simulations sequence. Panels (a)-(d) show the situation after 10, 30, 100 and 1000 iterations respectively. In this case, already 100, or even 30 simulations seem to cover the relevant portion of the parametric space

generic component θ_k can be estimated as:

$$\widehat{\text{Var}}(\theta_k \mid \mathbf{y}) = \frac{S-1}{S} W(\theta_k) + \frac{1}{S} B(\theta_k),$$

where $W(\theta_k)$ and $B(\theta_k)$ are the average within-chain variance and the

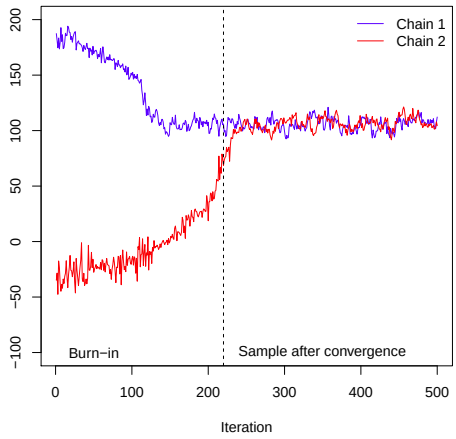


FIGURE 2.11
Graphical assessment of convergence for a Markov chain; in this case two chains are set up, starting from different initial points. After the burn-in period, the two chains converge to the stationary distribution

between-chains variance, and S is the length of the MCMC sample. Convergence is then monitored by assessing the *potential scale reduction*

$$\hat{R} = \sqrt{\frac{\widehat{\text{Var}}(\theta_k \mid \mathbf{y})}{W(\theta_k)}}, \tag{2.19}$$

which represents the factor by which the scale of the current estimated posterior distribution of θ_k can be further reduced. If $\hat{R}$ is large, then considering a longer MCMC run will potentially improve the inference about the target distribution. As a rule of thumb, values of $\hat{R} \leq 1.1$ are generally accepted as indicative of sufficient convergence.

2.4.7 MCMC autocorrelation

The second critical aspect of MCMC procedures is that the iterations produced by a Markov chain are by definition correlated, since the current observation depends on the previous one. Therefore, intuitively the actual number of iterations stored to produce the inference does not give in general the same information provided by a sample of *iid* observations of the same size. In other words, the higher the autocorrelation, the lower the degree of equivalence between the MCMC output and a proper *iid* sample of the same size.

3

Statistical cost-effectiveness analysis

3.1 Introduction

In the last ten years, health economic evaluations have built on more advanced statistical decision-theoretic foundations, effectively becoming a branch of applied statistics (Briggs et al., 2006; Willan and Briggs, 2006), increasingly often under a Bayesian statistical approach (O'Hagan and Stevens, 2001; O'Hagan et al., 2001; Parmigiani, 2002b; Spiegelhalter and Best, 2003; Spiegelhalter et al., 2004).

As suggested by Spiegelhalter et al. (2004), this can be ascribed to the fact that "the subjective interpretation of probability is essential, since the expressions of uncertainty required for a decision analysis can rarely be based purely on empirical data".

Even though the process is, technically, a simple application of standard decision-theoretic precepts (described for example in Lindley, 1985), health economics is complicated by issues related to other important factors that play a major role in real practice medical decision making. Among these are the difficulty of applying standard cost-effectiveness techniques to the regulatory process (Baio and Russo, 2009), and the necessity of properly accounting for the impact of uncertainty in the inputs of decision processes, an issue known as *sensitivity analysis* (Parmigiani, 2002b; Saltelli et al., 2004). This latter in particular is fundamental and is a required basic component of any new drug approval or reimbursement dossier in settings regulated by decision-making bodies such as NICE in the UK (Claxton et al., 2005).

In this chapter we first briefly review the main characteristics of decision theory. As in chapter 2 we proceed by introducing the more abstract theory, in order to make the point that rational decision-making is effected by maximising the expectation of a suitably defined utility function. This is used to quantify the value associated with the uncertain consequences of a possible intervention.

Next we link the general methodology to the specific problem faced in health economic evaluation. This requires the specification of the problem in terms of a composite response, accounting for both cost and benefits. We present a relatively simple running example and, as in chapter 2, we switch between the development of the theory and its application throughout.

We then concentrate on the development of sensitivity analysis techniques, which as suggested earlier play a fundamental role in health economic evaluations. Finally, we present some more advanced issues associated with the main assumptions on which cost-effectiveness or cost-utility analyses are based: in particular, we consider the problems of risk-aversion and the impact of market constraints (*e.g.* in the case of regulatory processes).

3.2 Decision theory and expected utility

3.2.1 The problem

Health economic evaluations are a typical problem of decision-making under uncertainty. The main objective is to evaluate comparatively the unknown consequences of a given health intervention against at least another. A suitable approach to deal with this kind of problems is based on *expected utility theory*, which we briefly review in this section. More substantial references are Savage (1954), Raiffa (1968), Lindley (1985), Berger (1985), Smith (1988), Bernardo and Smith (1999), Parmigiani (2002b), Jordaan (2005) and Smith (2011).

Formally, a decision problem is characterised by some fundamental elements: first, we consider the possible *decisions* (interventions, actions, treatments) $t \in \mathcal{T}$, representing the alternatives available to the decision-maker. The selection of each possible intervention has some *consequences* (outcomes) $o \in \mathcal{O}$, defined in general as functions of suitable random quantities $\boldsymbol{\omega} \in \Omega$. Every consequence can be expressed as $o = (\boldsymbol{\omega}, t)$, *i.e.* as the result of choosing intervention t and the fact that a series of random quantities $\boldsymbol{\omega}$ will obtain in the future. The set of consequences can be then represented as $\mathcal{O} = \Omega \times \mathcal{T}$.

In addition to these fundamental quantities, the decision-maker needs to define a scheme of *preferences* among the many decisions and consequences; this relationship of preference is generically indicated by the symbol '$\preceq$'. The notation $t_1 \preceq t_2$ indicates that the random consequences of action t_1 are *not* preferred to those of action t_2. If $t_1 \preceq t_2$ and simultaneously $t_2 \preceq t_1$, the two actions are indifferent: $t_1 \sim t_2$.

The Bayesian decision process is based on a set of *prescriptive axioms*. These are the criteria that *should* hold in order to make rational decisions. The first set of axioms is related to the *coherence* of the decision making and involves:

- *comparability of the consequences*. This assumes that the decision-maker is capable of producing some form of ranking of the possible outcomes, so that there exist at least one pair of consequences o_1 and o_2 for which the former is preferred to the latter;
- *transitivity of the preferences*. This axiom implies that if the decision-maker has a preference for action t_2 over action t_1 and for action t_3 over

n_{int} is the number of interventions being compared* (chapter 4 discusses how to run a Bayesian model, store and post-process its results using R).

In this case, $n_{\text{int}} = 2$ since we are comparing two interventions and we use $n_{\text{sim}} = 500$ simulations. The relevant health economic quantities can be produced by running the function bcea, by means of the following commands (cfr. §4.7 for a more detailed description).

```
treats <- c("Old Chemotherapy","New Chemotherapy")
m <- bcea(e=e,c=c,ref=2,interventions=treats,Kmax=50000)
```

The first command defines a vector of labels to be associated with each intervention. Then we create an object m which contains the results of the cost-effectiveness analysis performed by bcea.

This function takes several inputs: the most important ones are the two matrices e and c. Then we need to specify the reference intervention: in this case we set the option ref=2 which tells R that the second column of e and c contains the values simulated for the intervention being compared to the standard treatment. If the option ref is left unspecified, R assumes that the first intervention is to be used as the intervention under analysis.

Next we specify that the interventions have labels defined in the vector treats; if the option interventions was left unspecified, R would construct labels in the form "Intervention 1", ..., "Intervention T" (with T $= n_{\text{int}}$).

The last option is related to the willingness-to-pay parameter and specifies the maximum value to be used for the analysis. In this case, we set the option Kmax=50000, which instructs R to select a grid of values between 0 and 50 000. The grid is built by considering a point every Kmax/500 in that interval. If nothing is specified for Kmax, the function bcea automatically considers a value of 50 000.

All the results derived upon varying k in the grid are stored in the object m and the vector of values for k is saved in the element† m$k, which in this case contains the 501 values $(0, 500, 1000, \ldots, 50\,000)$.

Using the output produced by bcea and saved in m, we can simply obtain a graph of the cost-effectiveness plane by entering the command ceplane.plot(m,comparison=1,wtp=25000). The input of this function is the object m and there are two possible options: the first one specifies which comparison should be plotted. In this case, there are only two interventions and therefore there can only be one comparison. In general, there are $n_{\text{int}} - 1$

*Notice that, while specifically designed for a Bayesian analysis, BCEA can also be run in a frequentist setting, provided that the two matrices with simulations from the distributions of e and c are available. In a non-Bayesian setting, these might be obtained, for example, by using re-sampling algorithms such as the bootstrap (although, of course, in that instance they would not represent the posterior distributions).

†In R, the elements contained in an object can be accessed using the notation object$element. Thus, typing m$k prints the entire grid of values selected for the willingness-to-pay defined in the object m.

possible comparisons. The second option specifies the value of the willingness-to-pay to use as reference. In this case, we have chosen the default value of $k = 25\,000$, which is usually recommended by NICE as the reference cost-per-QALY.

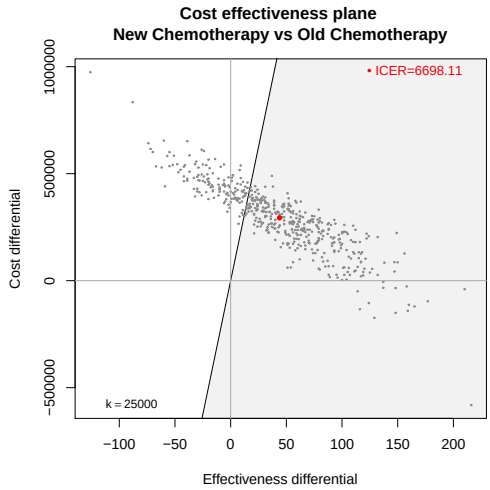


FIGURE 3.1
Cost-effectiveness plane for the chemotherapy example. The dots represent the simulations from the posterior distribution of (Δ_e, Δ_c), while the shaded part of the graph shows the "sustainability area", *i.e.* the portion of the plane in which the points are below the willingness-to-pay threshold, which is set to 25 000 in this case

The result is depicted in Figure 3.1, in which the dots are the simulations from the posterior distribution of (Δ_e, Δ_c). The graph also shows the line obtained in correspondence of the set value of k. The shaded area below the line represent the portion of the plane where the simulated values are below that threshold and therefore it can be considered as a "sustainability area".

The red dot represents the ICER (cfr. §1.7). As in this case it lies in the sustainability area, we can conclude that, at the willingness-to-pay threshold selected, the new drug is a cost-effective alternative with respect to the status quo. With respect to Figure 1.6, the current analysis also presents a quantification of the uncertainty underlying the point estimation represented by the ICER, because it is based on the entire distribution of (Δ_e, Δ_c), rather than just on its expectations.

A more comprehensive analysis is provided by Figure 3.2, which is produced

4

Bayesian analysis in practice

4.1 Introduction

As discussed in chapter 2, if it is possible to sample from the full conditional distributions, Gibbs sampling algorithms can be programmed in a relatively easy way. However, in most practical situations, the required conditional distributions are not analytically tractable and therefore it is necessary to approximate them (*e.g.* by means of algorithms such as Metropolis-Hastings or slice sampling) before Gibbs sampling can be performed.

The most popular software that allows the semi-automatisation of MCMC procedures is `BUGS`, and particularly its MS Windows incarnations `WinBUGS` (Spiegelhalter et al., 2002) and `OpenBUGS` (Lunn et al., 2009), whose widespread use has arguably contributed to the establishment of applied Bayesian statistics in the last twenty years.

The acronym `BUGS` stands for *Bayesian analysis Using Gibbs Sampling* and the program essentially consists of two main parts. The first is a *parser*, which inspects the set of declarations provided by the user to define the statistical model (in terms of data and parameter distributions and, possibly, other deterministic relationships among the variables in the problem). In particular, the parser codifies the statistical model in terms of the corresponding DAG, trying to make use of the conditional independence relationships implied by the model assumed by the user. These generally simplify the computations since the full conditional distribution for any (set of) node(s) only involves a local computation on the graph. Thus, only a small portion of the whole model needs to be considered at any given time (Lunn et al., 2009).

The second part is an expert system that is used to deduce the form of the full conditional distributions generated by the problem. When possible, `BUGS` tries to exploit conjugacy to speed up the process; when this is not feasible, suitable complementary sampling algorithms are applied together with the Gibbs sampling to obtain the required MCMC estimation.

While both `BUGS` and `WinBUGS` can run as stand-alone software, in recent years several programs have been written to interface them with standard statistical software such as `R`, `Matlab`, `Stata` or `SAS`, which makes the process of data analysis easier (we discuss this aspect later).

Despite their wide success, `WinBUGS` or `BUGS` are not the only possible al-

baby, knowing that the mother has gone into labour at exactly that estimated gestational age), or a hypothetical value (*e.g.* if the doctor wants to plan different care strategies for a mother who has not gone into labour yet).

We can then run the model using the following code.

```
X.star <- 28
dataJags2 <- list("N","y","X","k","X.star")
filein <- "modelNormal2.txt"
params2 <- c("alpha","beta","sigma","y.star")
inits <- function(){
   list(alpha=rnorm(1),beta=rnorm(1),
        lsigma=rnorm(1),y.star=runif(1,0,6000))
}
m2 <- jags(dataJags2, inits, params2, model.file=filein,
  n.chains=2,n.iter=50000,n.burnin=4500,n.thin=91,DIC=TRUE)
print(m2,digits=3,intervals=c(0.025, 0.975))
```

After we have set the value for the new estimated gestational age `X.star`, we redefine the data list to include this node as well. Then we redefine the name of the model file to point `JAGS` towards the new specification and the object containing the parameters to be monitored so that the predictive distribution of `y.star` is included.

The next modification to the previous `R` code is in the `inits` function, where we set an initial value also for the node `y.star`. Technically, this is not strictly necessary; in fact, `JAGS`/`BUGS` will estimate the predictive distribution effectively using a simple MC approach (such as the one discussed in §2.4.3) using for each iteration the current value of the relevant parameters and thus there is no issue of convergence. However, it is generally a good idea to provide reasonable starting values for any non-observed random quantity and in this case we do so by providing a value from a Uniform distribution in the interval $[0; 6000]$.

Finally, we run the `jags` function for 50000 iterations using the first 9500 for the burn-in and thinning of 81. This implies that the simulations saved to produce the posterior inference are 1000. The results are saved in the object `m2` which is then printed to give the following output.

```
Inference for Bugs model at "modelNormal2.txt", fit using jags,
 2 chains, each with 50000 iterations (first 9500 discarded),
 n.thin = 81, n.sims = 1000 iterations saved
           mu.vect sd.vect      2.5%     97.5%  Rhat n.eff
alpha    -2343.609 169.118 -2667.398 -2023.663 1.000  1000
beta       143.319   4.333   135.147   151.706 1.000  1000
sigma      455.764   9.928   436.599   476.001 1.002  1000
y.star    1677.155 460.291   743.366  2565.684 1.000  1000
deviance 16815.939   8.107 16802.916 16833.459 1.000  1000

For each parameter, n.eff is a crude measure of effective sample size,
and Rhat is the potential scale reduction factor
```

5

Health economic evaluation in practice

5.1 Introduction

In this final chapter we present some examples of health economic evaluation. In particular we focus on three "typical" cases; the first concerns the analysis of individual level data, specifically from a RCT, in which a sample of individuals is observed in terms of the relevant measures of cost and clinical outcome. The second example focusses on the process of evidence synthesis, a situation particularly relevant when individual data are not available. In these situations, the relevant random quantities can be estimated by the combination of the available evidence, *e.g.* coming from published studies, or expert opinions. Within the Bayesian framework, this is very much linked to the development of hierarchical models, which we briefly review before presenting the example. Finally, we consider the analysis of Markov models, an increasingly popular tool in health economic evaluation, which allow the simulation of a follow up analysis on a "virtual" cohort of patients.

While the problems highlighted in each of the following sections can be considered as typical of the situations considered in applied health economics, they are far from representing an exhaustive set: in real applications, there are countless subtleties and nuisances that need to be addressed specifically. In particular in the Bayesian approach, this require a careful specification of the model to be used, mainly in terms of the prior distributions, but also in terms of the possible correlation levels among the observed and unobserved random variables.

Nevertheless, we tackle some of the most relevant issues arising from the analysis of health economic data, trying to point out possible solutions and references where more detailed modelling strategies are presented. All the examples are worked out starting from the description of the problem, the specification of the Bayesian model and then the code used to run the MCMC analysis and the post-processing necessary to derive the relevant health economic quantities used to produce the decision-making process.

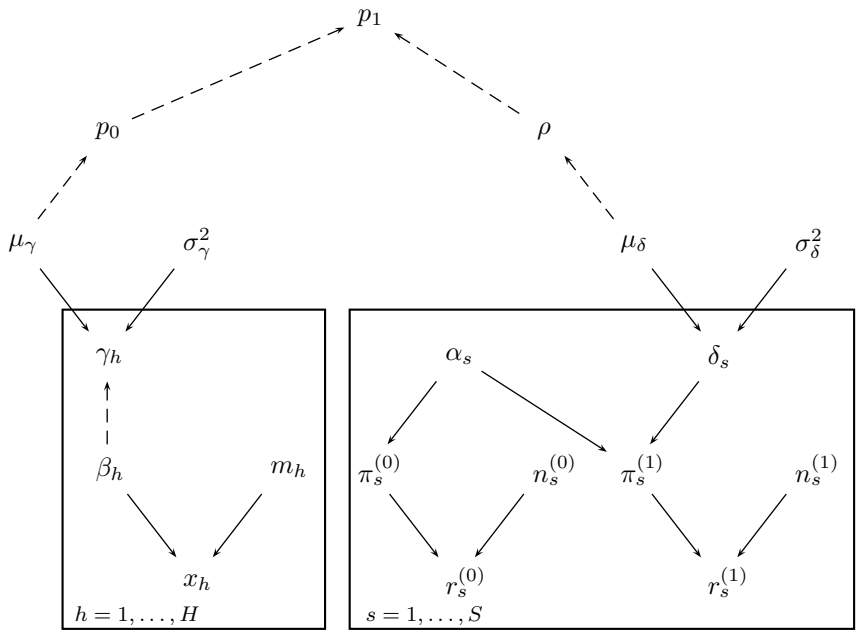


FIGURE 5.7
Graphical representation of the evidence syntheses in the model. In the graph, solid arrows indicate probabilistic links, while dashed arrows indicate logical dependence. H studies are used to investigate the overall population probability of being infected by influenza, p_0. A similar structure combines the information for the S studies investigating the effectiveness of NIs to derive an odds ratio, which is combined with the estimation of p_0 to provide an estimation of p_1, the probability of influenza in the scenario in which prophylactic treatment with NIs is made available

```
#   the "healthy" adults population (t=0)
    for(h in 1:H) {
        x[h] ~ dbin(beta[h], m[h])
        logit(beta[h]) <- gamma[h]
        gamma[h] ~ dnorm(mu.gamma,tau.gamma)
    }

# Evidence synthesis for effectiveness of NIs (t=1 vs t=0)
    for (s in 1:S) {
        r0[s] ~ dbin(pi0[s],n0[s])
        r1[s] ~ dbin(pi1[s],n1[s])
        logit(pi0[s]) <- alpha[s]
        logit(pi1[s]) <- alpha[s]+delta[s]
```

```
        delta[s] ~ dnorm(mu.delta,tau.delta)
        alpha[s] ~ dnorm(0,0.00001)
    }

# Prior distributions
    mu.delta ~ dnorm(0,0.00001)
    mu.gamma ~ dnorm(0,0.00001)
    sigma.delta ~ dunif(0,10)
    tau.delta <- pow(sigma.delta,-2)
    sigma.gamma ~ dunif(0,10)
    tau.gamma <- pow(sigma.gamma,-2)

# Costs of influenza
    c.inf ~ dnorm(mu.inf,tau.inf)
# Length of time to recovery when infected by influenza
    l ~ dlnorm(mu.l,tau.l)

# Odds Ratio of influenza under treatment with NIs
    rho <- exp(mu.delta)
# Estimated probability of influenza in "healthy adults" for t=0
    p0 <- exp(mu.gamma)/(1+exp(mu.gamma))
# Estimated probability of influenza in "healthy adults" for t=1
    p1 <- (rho*p0/(1-p0))/(1+rho*p0/(1-p0))
}
```

The data pre-processing required in R involves the definition of the variables containing the observed data. We do this with the following code.

```
# Evidence synthesis on incidence of influenza
#   in healthy adults (under t=0)
x <- m <- numeric()
x <- c(0,6,5,6,25,18,14,3,27)
m <- c(23,241,159,137,519,298,137,24,132)
H <- length(x)

# Evidence synthesis on effectiveness of NIs vs placebo
r0 <- r1 <- n0 <- n1 <- numeric()
r0 <- c(34,40,9,19,6,34)
r1 <- c(11,7,3,3,3,4)
n0 <- c(554,423,144,268,251,462)
n1 <- c(553,414,144,268,252,493)
S <- length(r0)

# Data on costs
unit.cost.drug <- 2.4      # unit (daily) cost of NI
length.treat <- 6*7        # 6 weeks course of treatment
```

Index