

# 多元贝叶斯回归模型框架下绘制成本-效果可接受曲线<sup>△</sup>

彭娟\*, 骆福添<sup>#</sup>(中山大学公共卫生学院, 广州 510080)

中图分类号 R956 文献标志码 C 文章编号 1001-0408(2013)10-0865-03

DOI 10.6039/j.issn.1001-0408.2013.10.01

**摘要** 目的:探讨成本-效果可接受曲线(CEAC)采用贝叶斯方法的原因,以及如何在贝叶斯回归模型框架下绘制CEAC。方法:首先从原理上分析贝叶斯方法的优势,其次在R软件中生成模拟数据并在Openbugs软件中实现模型,绘制CEAC,比较不同先验信息下该曲线的差异。结果:在强先验信息下,参数后验分布的均数比弱先验信息下更准确、标准差更小。在一定范围内,强先验信息下增量-净效益大于0的概率大于弱先验信息下的概率。结论:贝叶斯方法在参数估计、区间估计、概率解释上具有优势。在贝叶斯方法下,CEAC可得到合理的解释。贝叶斯方法先验信息对参数估计与成本-效果分析决策甚为重要。在强先验信息下,参数估计的准确度、变异性更小,同时CEAC也更准确。

**关键词** 贝叶斯回归;成本-效果;增量-净效益;模拟

## Cost-effectiveness Acceptability Curve Drawn from Multivariate Bayesian Regression Model Framework

PENG Juan, LUO Fu-tian (School of Public Health, Sun-Yat Sen University, Guangzhou 510080, China)

**ABSTRACT** **OBJECTIVE:** To investigate the reasons for cost-effectiveness acceptability curve (CEAC) using Bayesian methods, and how to draw CEAC in the framework of the Bayesian regression model. **METHODS:** Firstly, the advantages of Bayesian approach were analyzed from the principle. Secondly, R generated simulation data, Openbugs software implemented model and CEAC was drawn. The differences in the curves of different prior information were compared. **RESULTS:** Under strong prior information, the posterior distribution of the mean parameters was more accurate than weak prior information, and its standard deviation was smaller than weak prior information. The probability of incremental-net benefit which was greater than 0 under strong prior was greater than the probability of weak prior within limits. **CONCLUSIONS:** Bayesian method is characterized with the advantages of parameter estimation, interval estimation and probabilistic interpretation. By Bayesian approach, CEAC can provide a reasonable explanation. The Bayesian approach prior information is important for parameter estimation and cost-effectiveness analysis decision-making. Under strong prior information, the accuracy and variability of the parameter estimation become weak, while CEAC is more accurate.

**KEY WORDS** Bayesian regression; Cost-effectiveness; Incremental-net benefits; Stimulation

成本-效果估计不可避免地含有不确定性,包括来自成本的或效果的。与成本-效果的点估计和区间估计相比,成本-效果可接受曲线(CEAC)是一种直观地描绘成本-效果估计中不确定性的绘图方法。然而,传统方法在该曲线的解释上比较牵强,使得贝叶斯方法在成本-效果分析中得到广泛应用。尽管该方法在国外的研究和应用已相对较多,但在国内的研究却很少。在中国知网、万方数据库等搜索,只有几篇介绍该方法的文章。张玉哲等<sup>[1]</sup>介绍贝叶斯法可利用先验信息进行推测,可在理论上准确地提供增量成本-效果比、CEAC、净货币收益的概率阐述;吴晶等<sup>[2]</sup>给出了CEAC的计算方法,但如何实现CEAC,特别是在贝叶斯框架下如何实现CEAC并未讨论。因此,CEAC如何绘制、为什么要采用贝叶斯方法来绘制等问题亟待解决。本文就贝叶斯方法的优势、在贝叶斯回归框架下如何绘制CEAC作了深入探讨,同时比较了在不同先验

信息下CEAC的区别。文章分析基于模拟实验,模拟实验利用R与Openbugs软件实现。

### 1 贝叶斯方法的优势

从贝叶斯估计来讲,在获得样本之前,有一部分经验的、历史的信息称为先验信息。贝叶斯参数的后验分布其实就是似然函数和先验分布的乘积。而贝叶斯估计是通过后验分布得到的。因此,在无先验信息或弱先验信息(参数几乎等概率的在取值范围内取值)下,贝叶斯估计应该与极大似然估计一致。当先验分布有效利用了先验信息时,贝叶斯估计应该更加准确,抽样随机性影响更小。从贝叶斯区间估计来讲,贝叶斯方法假设参数是随机的,可计算出该参数落入任一子域内的概率;对给定概率,可找到一些区域,贝叶斯区间选择使得该区域上的后验密度值不少于相同概率下其他区域上的值,也就是最高后验密度的区域即相同概率下最短的区域。由此得到的区间可解释为该参数落入其中的概率为 $1-\alpha$ ,这与实

## 本栏目协办

南京正大天晴制药有限公司

地址:江苏省南京市玄武区长江路188号德基大厦22层  
电话:025-86816983 邮编:210018

<sup>△</sup>基金项目:卫生部科技专项课题资助项目(No.2008ZX09312002-003)

\* 硕士研究生。研究方向:成本-效果分析中贝叶斯方法的应用。E-mail:dielianshuayaya@163.com

<sup>#</sup> 通信作者:教授,博士研究生导师,博士。研究方向:统计方法。电话:020-87330730。E-mail:luoft@sysu.edu.cn

际中的理解吻合。而传统方法通过构造统计量,解释为当抽样次数达到很大时,大约有 $100 \times (1-\alpha)$ 个区间覆盖真值。 $1-\alpha$ 为置信水平,在解释上比较牵强。因此,贝叶斯方法综合利用了先验信息和样本信息,对统计推断能给出一个直观的解释。而传统方法作出的推断总是基于频率的基础,对估计值的统计性质实际上是对大量重复实验作出的推断,不能对当前估计值的统计性质给出一个明确的答案,在应用中含有更多的不确定性。

## 2 探讨贝叶斯框架下CEAC的计算

当卫生决策者在决定是否补偿某种卫生技术时,比较新技术和已有技术的效果和成本-效果分析显得越来越重要。假设在一个临床试验中,要比较2种医疗技术 $T_1, T_2$ 。数据

$$D = \{x_{ij} = (e_{ij}, c_{ij})^T; i=1, 2, j=1, 2, \dots, n_i\}$$

$e_{ij}, c_{ij}$ 分别代表接受第*i*种治疗的第*j*个患者的效果和成本。为了比较哪一个方案的成本-效果更好,首先得知道各个方案的期望成本和期望效果。令 $\Delta E = \mu_2 - \mu_1, \Delta C = \gamma_2 - \gamma_1$ ,代表平均效果差异和平均成本差异。常用指标是增量成本-效果比 $\rho = \Delta C / \Delta E$ ,代表在原有基础上额外增加1个单位效果所需的额外成本。当比值 $\rho > \lambda$ ,就代表 $T_2$ 是更好的。 $\lambda$ 表示社会为增加单位效果支付的最大数额,此时 $\lambda$ 是某个固定的阈值。

另一方面, $\rho > \lambda$ 等价的表示为增量净效益(INB)  $> 0$ 。 $INB = \lambda \Delta E - \Delta C$ , $\lambda$ 是为增加单位效果愿意支付的最大货币值,是可变的常数。同样,由于 $\Delta E, \Delta C$ 估计的不确定性,INB的估计值也具有不确定性,并且INB估计的分布由 $\Delta E, \Delta C$ 的联合分布来表示。当 $\lambda$ 取值固定时,可得到 $INB > 0$ 的概率。对一系列 $\lambda$ ,可绘制 $INB > 0$ 的概率的曲线图,即CEAC。原理如下:

考虑第*i*种治疗的抽样模型 $f(e_i, c_i | \theta_i), i=1, 2, \theta_i$ 代表未知的不可观察的参数。数据的似然函数为:

$$l(\theta_1, \theta_2 | data) = \prod_{i=1}^2 \prod_{j=1}^{n_i} f(e_{ij}, c_{ij} | \theta_i) \dots \dots \dots (1)$$

治疗方案间独立,对某个贝叶斯模型, $\pi(\theta_i), i=1, 2$ 为 $\theta_i$ 的先验分布。假设参数间是独立的, $(\theta_1, \theta_2)$ 联合后验分布为:

$$\pi(\theta_1, \theta_2 | data) = \frac{l(\theta_1, \theta_2 | data) \pi(\theta_1) \pi(\theta_2)}{\int \int l(\theta_1, \theta_2 | data) \pi(\theta_1) \pi(\theta_2) d\theta_1 d\theta_2} \dots (2)$$

临床试验中,期望用个体的临床特点和接受的治疗方案的线性组合来解释患者的成本和效果,于是多重贝叶斯回归可写成如下形式:

$$\begin{aligned} E_i &= b_0 + b_1 x_{1,i} + b_2 x_{2,i} + \dots + b_{k-1} x_{k-1,i} + b_r x_{r,i} + e_i \\ C_i &= c_0 + c_1 x_{1,i} + c_2 x_{2,i} + \dots + c_{k-1} x_{k-1,i} + c_r x_{r,i} + v_i \\ \begin{pmatrix} e_i \\ v_i \end{pmatrix} &\sim N_2 \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma \right), \lambda = \Sigma^{-1} = \begin{pmatrix} \sigma_e^2 & \sigma_{ec} \\ \sigma_{ec} & \sigma_c^2 \end{pmatrix}^{-1} \dots \dots \dots (3) \end{aligned}$$

假设 $(e_i, v_i)$ 是误差项,是独立的且一致服从多元正态分布,为精度矩阵。指定参数服从多元正态先验分布,即假设 $b \sim mvn(\beta^0, v_1^{-1}), c \sim mvn(c^0, v_2^{-1})$ ,对协方差结构服从逆威沙特分布 $\lambda \sim IW(A, f)$ <sup>[3]</sup>。

变量 $x_n$ 的系数 $b_r, c_r$ ,分别代表新、旧治疗方案相比平均增加的效果和成本,其他协变量的系数也类似解释,于是增量成本-效果的联合分布体现为 $b_r, c_r$ 的联合分布。在马尔可夫链蒙特卡洛模拟技术(MCMC)算法下,得到一系列来自参数联合后验分布的样本。绘制CEAC得计算不同 $\lambda$ 下, $INB > 0$ 的概率。而INB的分布较难获得,可通过来自INB的分布的样本来估计概率。

由 $INB = \lambda \Delta E - \Delta C$ ,而 $\Delta E, \Delta C$ 的样本可通过MCMC方法

得到,因此可直接利用回归系数的模拟样本得到INB的样本。在大样本的基础上,可利用大数定律直接得到 $INB > 0$ 的概率,从而绘制出CEAC。具体实现见程序部分。

## 3 模拟实验分析

在R软件中生成模拟数据,模型程序在Openbugs软件中实现,所用模拟数据来自模型:

$$\begin{aligned} effect_i &= 20 + 1 * age_i + 10 * treat_i + \varepsilon_{1i} \\ cost_i &= 100 + 2 * age_i + 30 * treat_i + \varepsilon_{2i} \end{aligned}$$

$$\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \sim dmnorm(0_{2 \times 2}, \Sigma), \Sigma = \begin{pmatrix} 10 & 0 \\ 0 & 100 \end{pmatrix} \dots \dots \dots (4)$$

用模型(4)生成模拟数据,在Openbugs中实现多元贝叶斯回归模型,模型的结果包括参数估计以及不同 $\lambda$ 下 $INB > 0$ 的概率。主体程序如下:

```
model{
  for (i in 1:N){
    output[i, 1:2] ~ dmnorm(mu[i, 1:2], v[1:2, 1:2]);
#output = [cost, effect]
    mu[i, 1] <- alpha[1] + alpha[2]*age[i] + alpha[3]*treat[i];
    mu[i, 2] <- beta[1] + beta[2]*age[i] + beta[3]*treat[i];
  }
  v[1:2, 1:2] ~ dwish(A1[, ], f1); #精度阵服从威沙特分布
  alpha[1:3] ~ dmnorm(a[, ], A[, ]); #系数服从多元正态分布
  beta[1:3] ~ dmnorm(b[, ], B[, ]);
  for (k in 1:NK) {
    Q[k] <- step(K[k]*alpha[3] - beta[3]) #Q[k] = 1 如果式子 > 0
  }
  var[1:2, 1:2] <- inverse(v[1:2, 1:2])
  deviation1 <- sqrt(var[1, 1])
  deviation2 <- sqrt(var[2, 2])
  correlation <- var[1, 2] / (deviation1 * deviation2)
}
```

模型的先验分布用真值 $a=c(20, 1, 10), b=c(100, 2, 30)$ 作为均数。设定A, B的值作为可以表示对系数alpha, beta估计值的把握度。如在强先验信息下,令上述均数的变异度为0.01;在弱先验信息下,令上述均数的变异度为100。此时,a, b可取范围内的任意值,相当于没有特定先验信息。显然,在强先验信息的情况下,参数估计的变异程度和准确度应该更好。如表1所示,各个参数第1行是弱先验信息的估计,第2行是强先验信息的估计。实际情况与预期吻合。不管是强先验信息,还是弱先验信息,MC\_error都很小,说明本次模拟结果很稳定。从参数后验分布的均值和标准差都可看出强先验信息下的估计准确度高,变异程度小;中位数和均数比较差异均 $< 0.01$ ,说明参数后验分布对称,符合正态分布假设。

图1以强弱先验信息下绘制的CEAC为例来说明2种方法在参数解释上的差异。INB( $R_c$ ) =  $R_c \Delta E - \Delta C$ ,  $R_c$ 表示增量效果愿意支付的货币值,INB代表增量净效益,prob代表 $INB > 0$ 的概率。

图1中,曲线在点 $R_c = 0.93$ 时, $p = 0.50$ ,贝叶斯派的解释为当增加单位效果患者愿意付超过0.93元时,选择治疗有较大的把握能够获得盈利,而不是亏本。这里的把握程度会随着 $R_c$ 的取值慢慢增大,即概率值增大。但是按频率派的解释,当患者愿意支付超过0.93元时,重复100次这样的治疗,大约有超过一半的治疗会盈利,而不是亏本,至于该次治疗的增量-净

表1 强弱先验下参数估计比较

Tab 1 Comparison of parameter estimation under strong and weak prior information

参数	mean	sd	MC_error	val2.5pc	median	val97.5pc	
效果项	常数	29.490 0	7.371 0	0.098 1	14.940 0	29.450 0	44.070 0
		20.000 0	0.032 7	0.000 4	19.940 0	20.000 0	20.060 0
	年龄	0.775 9	0.246 6	0.003 3	0.291 3	0.771 0	1.258 0
		0.996 7	0.028 0	0.000 3	0.941 6	0.996 4	1.052 0
	治疗方案	3.826 0	3.244 0	0.043 0	-2.461 0	3.835 0	10.310 0
		9.999 0	0.036 4	0.000 4	9.938 0	10.000 0	10.060 0
成本项	常数	136.400 0	7.443 0	0.091 1	121.600 0	136.500 0	150.900 0
		100.000 0	0.099 4	0.001 1	99.810 0	100.000 0	100.200 0
	年龄	1.242 0	0.247 2	0.003 0	0.772 8	1.237 0	1.734 0
		1.984 0	0.051 8	0.000 6	1.884 0	1.983 0	2.087 0
	治疗方案	2.329 0	3.334 0	0.040 1	-4.248 0	2.328 0	8.911 0
		29.980 0	0.100 2	0.001 2	29.780 0	29.980 0	30.180 0

注: MC\_error: 蒙特卡罗模拟误差; val2.5pc: 2.5%分位数; val97.5pc: 97.5%分位数

note: MC\_error: Monte Carlo simulation error; val2.5pc: 2.5% quantile; val97.5pc: 97.5% quantile

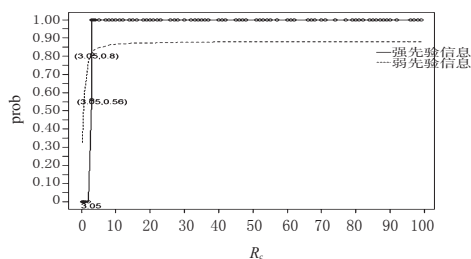


图1 成本-效果可接受曲线

Fig 1 Cost-effectiveness acceptability curves

效益不会  $> 0$ , 是未知的, 无法作出推断。

从图1中可看出, 在  $R_c = 3.05$  时, 强先验信息下  $INB > 0$  的概率为 0.56, 而弱先验信息下的概率达 0.8。在强先验信息下, 选择第 2 种方案的把握不是很大, 有可能选择第 1 种方案; 而在弱先验信息下, 毫无疑问选择第 2 种方案。说明在意愿支付额较小时, 决策的不确定性较大。弱先验信息下的结论并没能反映出此时决策的不确定性。而在  $R_c > 3.05$  时, 强先验信息下, 对本次数据结果的  $INB > 0$  的概率均大于弱先验信息下的概率。表明在有可靠先验信息的情况下, 对同一数据的结果所得  $INB > 0$  的把握度(信念)是不一样的。进一步说明, 在信息准确的情况下, 先验信息可增强结果的可靠性。

## 4 结果与讨论

在意愿支付法下, 本文引出了增量-净效益, 进而自然提出 CEAC。然而, 对 CEAC 的解释有学者提出质疑, 认为在贝叶斯框架下解释才合理。有必要指出的是, 本文中增量-净效益有文现状也称净货币效益。Briggs AH 等<sup>[4]</sup>定义  $INB(R_c) = R_c \Delta E - \Delta C$ 。  $R_c$  是用来度量单位效果差异的货币。Löthgren M 等<sup>[5]</sup>同样用该式来定义 INB。两种说法相同。

本文详细地分析了成本-效果回归框架下 CEAC 是如何绘制的问题, 并通过实际数据分析给出了程序。同时, 通过 CEAC 的比较, 可得出当前支付额度下, 某种治疗方案或技术成本效果更优的概率有多大。本文研究显示, 对不同的先验信息、经验信息越准确, 把握度越大, 在意愿支付法下, 对当前决策的把握会更大。

对 CEAC 的计算, 有学者<sup>[6]</sup>提出 INB 基于平均值的计算优于对称分布, 对偏态分布的效果较差, 因此有学者提出另一种计算方法——预测联合密度方法。当然, 本文对成本-效果分析中的一些问题讨论得还不够全面, 如在贝叶斯回归框架下, CEAC 的推理、MCMC 方法的具体计算等。另外, 有学者将研究扩展到对 3 种以上治疗技术的比较<sup>[7]</sup>。

## 参考文献

- [1] 张玉哲, 吴晶, 孙利华. 药物经济学评价中的贝叶斯法介绍[J]. 中国药物经济学, 2009(2): 51.
- [2] 吴晶, 刘国恩. 成本-效果可接受曲线: 不确定状态下的医疗决策方法[J]. 中国药物经济学, 2006(3): 52.
- [3] 朱慧明, 林静. 贝叶斯计量经济模型[M]. 北京: 科学出版社, 2009: 66-67.
- [4] Briggs AH. A bayesian approach to stochastic cost-effectiveness analysis [J]. *Health Econ*, 1999, 8(3): 257.
- [5] Löthgren M, Zethraeus N. Definition, interpretation and calculation of cost-effectiveness acceptability curves [J]. *Health Econ*, 2000, 9(7): 623.
- [6] Hernández MA, Vázquez-Polo FJ, González-Torre FJ, et al. Complementing the net benefit approach: a new framework for bayesian cost-effectiveness analysis[J]. *Int J Technol Assess Health Care*, 2009, 25(4): 537.
- [7] 宗欣, 孙利华. 成本-效果可支付曲线的理论与应用[J]. 中国药房, 2012, 23(10): 867.

(收稿日期: 2012-07-30 修回日期: 2012-10-24)

## 卫生部部长陈竺出席“畅想未来, 百年健康”会议

本刊讯 2013年1月26日, 卫生部部长陈竺在人民大会堂出席了协和医学院和洛克菲勒基金会共同举办的“畅想未来, 百年健康”会议暨洛克菲勒基金会百年纪念活动。

开幕式上, 陈竺部长简要回顾了协和医学院与洛克菲勒基金会合作的历史以及合作对中国医疗卫生做出的贡献, 感谢基金会对中国的帮助和支持。陈竺希望洛克菲勒基金会作为最早参与全球卫生并具有世界影响的基金会之一, 能够把握未来发展趋势, 在未来 100 年对提升世界人民福祉做出更大贡献。

洛克菲勒基金会成立于 1913 年, 1917 年资助了中国现代医学的摇篮——北京协和医学院, 并在 20 世纪 30 年代支持了中国农村卫生保健模式的创建。

来自世界各国的政府官员、知名专家、医学院院长、商界和民间社会领袖等出席了会议。会议就今后 25 年、50 年乃至 100 年的健康发展趋势进行了探讨, 并讨论了下个世纪建设健康社会的途径。

卫生部国际司、协和医学院有关负责人出席了会议。