

文章编号:1003-207(2009)04-0170-08

基于多子样的贝叶斯动态过程能力估计与评价方法研究

朱慧明¹, 曾惠芳¹, 虞克明², 郝立亚¹, 李素芳¹

(1. 湖南大学工商管理学院, 湖南 长沙 410082; 2. 布鲁内尔大学数学系, 伦敦 UB8 3PH)

摘要:针对参数随机化情况下生产过程能力的评价问题,提出了新的过程能力指数估计与评价方法。通过质量控制模型的统计结构分析,研究了扩散先验分布下参数后验分布,据此构造了过程能力指数的贝叶斯点估计和区间估计;在此基础上,将前一阶段模型参数后验分布作为下一阶段的参数先验分布,充分利用历史数据信息,建立了过程能力指数及其下限的贝叶斯动态评价模型。研究结果表明:与现有的贝叶斯过程能力指数估计方法比较,贝叶斯动态过程能力指数的预测精度优于前者,更能反映实际生产过程能力水平。

关键词:质量控制;过程能力指数;贝叶斯方法;估计;先验分布

中图分类号:C931.1 **文献标识码:**A

1 引言

统计过程控制是现代化工业生产中确保产品质量和可靠性的核心方法之一,它不仅在制造业中获得了广泛应用,而且也可以用于物流采购等非制造行业的质量评价问题^[1];而过程能力指数(Process Capability Indices,简称 PCI)则是度量一个工业过程能力的重要指标。过程能力就是生产过程处于统计控制状态下产品的加工质量满足技术标准的能力,过程能力指数是指过程能力满足产品质量标准要求(规格范围等)的程度。作为度量过程能力是否满足规格要求的过程能力指数,在制造业中获得了广泛应用。最早的过程能力指数是美国学者 Juran (1974)^[2]于1974年提出来的 C_p 指数,除了 C_p 过程能力指数外,还有许多其它描述过程能力的指数,例如,Chan, Cheng 和 Spiring (1988)^[3]的 C_{pm} 指数, Gupta 和 Kotz (1997)^[4]的 C_{pq} 指数,田口过程能力指数^[5],以及多变量过程能力指数^[6]等。同时,许多学者从不同的角度对过程能力指数进行了改进,如文献[7-10]的非正态分布的过程能力指数,文献[11]关于过程能力的 Bootstrap 抽样估计。

但是,现有的过程能力指数估计与评价方法均建立在经典统计(即频率统计)理论基础之上,其基本假设之一是模型参数不具有随机性,它没有考虑参数不确定性这一风险因素。实际上,生产制造环境是随时间的推延而变化的,产品质量性能指标的分布参数具有随机性的特征,这与经典统计的基本观点“总体分布参数是固定的常数”是不相符的。Lorenco (1998)^[12]和 Bernardo (1996)^[13]等学者认为:根据贝叶斯方法研究过程能力分析等质量管理问题更符合实际情况,因为在贝叶斯理论体系中,模型参数是随机变量,具有随机性。随着计算技术的发展,贝叶斯方法质量管理领域获得了广泛应用。在贝叶斯理论体系中,模型参数后验分布比较复杂,并不属于目前已有的分布范畴,造成贝叶斯方法难以在实际中获得应用。20世纪90年代以后,马尔科夫链蒙特卡罗仿真方法应用于贝叶斯估计问题,解决了贝叶斯计算瓶颈问题^[14]。目前,贝叶斯方法在质量控制和统计过程能力评价方面中应用文献不断增加^[15]。例如, Cheng 和 Spiring (1989)^[16], Shiao, Chiang 和 Hung (1999)^[17], Shiao, Hung 和 Chiang (1999)^[18], Wu (2008)^[19]利用多个多样数据分析了过程能力指数的贝叶斯估计, Singh 和 Saxena (2005)^[20]研究了过程能力指数的贝叶斯 shrinkage 估计问题, Bernardo 和 Irony (1996)^[21], 朱 (2005)^[22], 以及 Nivertthi 和 Dey (2000)^[23]讨论了多变量过程能力指数的贝叶斯估计。这些研究成果的

收稿日期:2008-12-22;修订日期:2009-04-20

基金项目:国家自然科学基金项目(70770138);教育部新世纪人才支持计划项目(NCET050704)

作者简介:朱慧明(1966-),男(汉族),湖南大学工商管理学院,教授、博士生导师,研究方向:贝叶斯预测与决策分析。

出现极大地促进了贝叶斯质量管理理论的发展,但它们属于静态贝叶斯分析方法范畴,难以融合产品历史数据信息,同时,也存在先验分布超参数确定的难题问题。

本文拟采用贝叶斯动态估计方法研究过程能力问题,在初始阶段选择使熵达到最大值的扩散先验分布(diffuse prior)作为模型参数的先验分布,从而不需要考虑超参数的取值问题,并将过程能力充足条件下的参数后验分布作为下一分析阶段参数的先验分布,充分利用历史数据信息分析问题,体现了贝叶斯理论的核心思想。本文的目的在于建立过程能力指数估计与评价的贝叶斯分析方法。

2 贝叶斯评价模型

2.1 贝叶斯估计

假设变量 X 为所需要考虑的产品质量特性指标,其统计模型如下:

$$X = \mu + \epsilon \tag{1}$$

此处 μ 为总体期望, ϵ 为随机误差项,它服从均值为零,方差为 σ^2 的正态分布,即 $\epsilon \sim N(0, \sigma^2)$, $\sigma^2 > 0$,其概率分布密度函数为

$$f(x / \mu, \sigma^2) = \left(\frac{1}{2\sigma^2}\right)^{1/2} \exp\left\{-\frac{1}{2\sigma^2}(x - \mu)^2\right\} \tag{2}$$

如果在过程能力分析的初始阶段,随机抽取了 m_1 批产品,获得 X 的相应样本观察值,结果如下:

$$x_i^{(1)} = \{x_{i1}^{(1)}, x_{i2}^{(1)}, \dots, x_{m_i}^{(1)}\}, i = 1, 2, \dots, m_1 \tag{3}$$

为方便起见,记

$$X^{(1)} = \{x_1^{(1)}, x_2^{(1)}, \dots, x_{n_k}^{(1)}\}, n_1 = \sum_{k=1}^{m_1} n_k^{(1)} \tag{4}$$

此时,对于给定的样本 $X^{(1)}$,模型参数 μ 和 σ^2 的似然函数为

$$\begin{aligned} L_1(\mu, \sigma^2 / X^{(1)}) &= \prod_{i=1}^{m_1} \prod_{j=1}^{n_i^{(1)}} f(x_{ij}^{(1)} / \mu, \sigma^2) \\ &= \prod_{i=1}^{m_1} \prod_{j=1}^{n_i^{(1)}} \left(\frac{1}{2\sigma^2}\right)^{1/2} \exp\left\{-\frac{1}{2\sigma^2}(x_{ij}^{(1)} - \mu)^2\right\} \\ &= \left(\frac{1}{2\sigma^2}\right)^{n_1/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{m_1} \sum_{j=1}^{n_i^{(1)}} (x_{ij}^{(1)} - \mu)^2\right\} \end{aligned} \tag{5}$$

为了进行贝叶斯分析,需要确定模型参数的先验分布。一般说来,在过程能力评价的初始阶段,产品的历史数据信息相对较少,根据我们选择使熵达

到最大值的扩散先验分布(reference prior)作为 μ 和 σ^2 的先验分布,即

$$p_1(\mu, \sigma^2) = \frac{1}{\sigma^2}, -\infty < \mu < \infty, \sigma^2 > 0 \tag{6}$$

根据贝叶斯定理,参数 (μ, σ^2) 的联合后验分布密度函数与其先验分布密度函数和似然函数的乘积成正比关系,即

$$p_1(\mu, \sigma^2 / X^{(1)}) \propto \left(\frac{1}{2\sigma^2}\right)^{(n_1+2)/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{m_1} \sum_{j=1}^{n_i^{(1)}} (x_{ij}^{(1)} - \mu)^2\right\} \tag{7}$$

由 $p_1(\mu, \sigma^2 / X^{(1)})$ 在区间 $(0, \infty)$ 上对 σ^2 进行积分计算,得参数 μ 的后验边缘分布:

$$\begin{aligned} p_1(\mu / X^{(1)}) &= \int_0^\infty p_1(\mu, \sigma^2 / X^{(1)}) d\sigma^2 \\ &= \int_0^\infty \left(\frac{1}{2\sigma^2}\right)^{(n_1+2)/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{m_1} \sum_{j=1}^{n_i^{(1)}} (x_{ij}^{(1)} - \mu)^2\right\} d\sigma^2 \\ &= \frac{1}{(\mu - \bar{x}^{(1)})^2 + \tilde{v}_1/m_1} \tag{8} \end{aligned}$$

其中

$$\tilde{v}_1 = \sum_{i=1}^{m_1} \sum_{j=1}^{n_i^{(1)}} (x_{ij}^{(1)} - \bar{x})^2, \bar{x} = \frac{1}{n_1} \sum_{i=1}^{m_1} \sum_{j=1}^{n_i^{(1)}} x_{ij}^{(1)} \tag{9}$$

显而易见,对于给定的样本 $x^{(1)}$, μ 的后验分布是自由度 $m_1 - 1$,位置参数 $\bar{x}^{(1)}$,尺度参数 \tilde{v}_1/m_1 ($m_1 - 1$) 的分布,即 $\mu / x^{(1)} \sim t(m_1 - 1; \bar{x}^{(1)}, \tilde{v}_1/m_1 (m_1 - 1))$;因此, μ 的后验期望等于其位置参数: $E(\mu / x^{(1)}) = \bar{x}^{(1)}$;在二次损失函数的条件下, μ 的贝叶斯点估计为其后验期望,所以 $\mu_B = \bar{x}^{(1)}$ 。

同样地,由 $p_1(\mu, \sigma^2 / X^{(1)})$ 在区间 $(-\infty, \infty)$ 上对 μ 积分计算,得参数 σ^2 的后验边缘分布:

$$\begin{aligned} p_1(\sigma^2 / X^{(1)}) &= \int_{-\infty}^{\infty} p_1(\mu, \sigma^2 / X^{(1)}) d\mu \\ &= \left(\frac{1}{2}\right)^{(n_1+2)/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{m_1} \sum_{j=1}^{n_i^{(1)}} (x_{ij}^{(1)} - \mu)^2\right\} d\mu \\ &= \left(\frac{1}{2}\right)^{(n_1+1)/2} \exp\left\{-\frac{1}{\sigma^2}\right\} \end{aligned} \tag{10}$$

此处 $\nu_1 = 2/\tilde{v}_1$ 。

由于

$$\int_0^\infty p_1(\sigma^2 / X^{(1)}) d\sigma^2 = 1 \tag{11}$$

而

$$\int_0^\infty \left(\frac{1}{2}\right)^{(n_1+1)/2} \exp\left\{-\frac{1}{\sigma^2}\right\} d\sigma^2 =$$

$$\left(\frac{m-1}{2}\right) \binom{n_1-1}{1}^{(n_1-1)/2} \quad (12)$$

因此, $\hat{\sigma}^2$ 的后验边缘分布密度函数为

$$f\left(\frac{\sigma^2}{X^{(1)}}\right) = \frac{1}{\left(\frac{m-1}{2}\right) \binom{n_1-1}{1}^{(n_1-1)/2}} \exp\left(-\frac{1}{\sigma^2}\right), \quad \sigma^2 > 0 \quad (13)$$

若记 $\lambda = 1/\sigma^2$, $\nu = (m-1)/2$, 根据随机变量分布密度和随机变量函数分布密度之间的函数关系, 可以证明: λ 的分布密度函数如下:

$$f(\lambda / X^{(1)}) = \frac{1}{\binom{\nu-1}{1} \lambda^\nu} \exp\left(-\frac{\lambda}{X^{(1)}}\right), \quad \lambda > 0 \quad (14)$$

不难看出, 对于给定的样本 $X^{(1)}$, λ 的后验边缘分布是形状参数为 ν , 尺度参数为 $1/X^{(1)}$ 的伽玛分布, 即 $\lambda / X^{(1)} \sim (\nu, 1/X^{(1)})$, 其后验期望为 $E(\lambda / X^{(1)}) = \nu / X^{(1)}$ 。根据 λ 的后验分布, $\sigma^{1/2}$ 的后验期望如下:

$$E(\sigma^{1/2} / X^{(1)}) = \int_0^{+\infty} \sigma^{1/2} f(\lambda / X^{(1)}) d\lambda = \int_0^{+\infty} \frac{1}{\binom{\nu-1}{1} \lambda^\nu} \exp\left(-\frac{\lambda}{X^{(1)}}\right) d\lambda = \frac{(\nu+0.5)}{\binom{\nu-1}{1} \lambda^{1/2}} \quad (15)$$

根据 Juran 的 C_p 过程能力指数的定义, 它是公差范围的长度与容忍区间长度之比, 即

$$C_p = \frac{USL - LSL}{6\sigma} \quad (16)$$

此处 USL, LSL 为上下公差限, σ 为过程标准差。据此计算过程能力指数 C_p 的后验期望

$$E(C_p / X^{(1)}) = E\left(\frac{USL - LSL}{6\sigma^{1/2}} / X^{(1)}\right) = \frac{(\nu+0.5)(USL - LSL)}{6 \binom{\nu-1}{1} \lambda^{1/2}} \quad (17)$$

和后验方差

$$Var(C_p / X^{(1)}) = E[(C_p)^2 / X^{(1)}] - [E(C_p / X^{(1)})]^2 = \frac{(USL - LSL)^2}{36 \binom{\nu-1}{1}} \left[1 - \left(\frac{\nu+0.5}{\binom{\nu-1}{1}}\right)^2\right] \quad (18)$$

因此, 在二次损失函数下, 过程能力指数 C_p 的贝叶斯点估计为

$$\hat{C}_{pB}^{(1)} = \frac{(\nu+0.5)(USL - LSL)}{6 \binom{\nu-1}{1} \lambda^{1/2}} \quad (19)$$

进一步, 设 $G(\nu, 1)$ 为伽玛分布 $(\nu, 1)$ 的 $(0 < \alpha < 1)$ 分位数, 则

$$P(G(\nu, 1) / X^{(1)}) = \quad (20)$$

即

$$P(\sigma > G(\nu, 1) / X^{(1)}) = 1 - \quad (21)$$

另一方面,

$$\begin{aligned} & P\left\{\sigma > G(\nu, 1) / X^{(1)}\right\} \\ &= P\left\{\sigma^{1/2} > \sqrt{G(\nu, 1)} / X^{(1)}\right\} \\ &= P\left\{\frac{USL - LSL}{6\sigma^{1/2}} > \frac{USL - LSL}{6\sqrt{G(\nu, 1)} / X^{(1)}}\right\} \\ &= P\left\{\frac{USL - LSL}{6} > \frac{USL - LSL}{6} \sqrt{G(\nu, 1)} / X^{(1)}\right\} \\ &= P\left\{C_p > \sqrt{G(\nu, 1)} / X^{(1)}\right\} \quad (22) \end{aligned}$$

因此

$$P\left\{C_p > \frac{USL - LSL}{6} \sqrt{G(\nu, 1)} / X^{(1)}\right\} = 1 - \quad (23)$$

从而, 过程能力指数 C_p 下限的贝叶斯估计为

$$\hat{C}_{pB} = \frac{USL - LSL}{6} \sqrt{G(\nu, 1)} \quad (24)$$

也就是说, C_p 的置信水平 $1 - \alpha$ 的单侧置信区间估计为 $\left[\frac{USL - LSL}{6} \sqrt{G(\nu, 1)}, \infty\right)$ 。

2.2 过程能力的评价

过程能力的评价过程包括以下三个步骤:

(1) 过程能力涵义的界定。一般说来, 在置信水平 $1 - \alpha$ 为 0.95 的情况下, 如果过程能力指数的取值大小不低于 1.33, 则可以认为过程能力充足。

(2) 过程能力的估计。先计算 ν , λ 和 $G(\nu, 1)$ 的取值, 然后根据 (19) 和 (24) 计算过程能力指数的贝叶斯点估计 $\hat{C}_{pB}^{(1)}$ 和过程能力指数下限的贝叶斯估计 $\hat{C}_{pB}^{(1)}$, 并确定其贝叶斯单侧置信区间。

(3) 过程能力的评价。显然, 过程能力指数的值越大, 表明产品的离散程度相对于技术标准的公差范围越小, 因而过程能力就越高; 过程能力指数的值越小, 表明产品的离散程度相对公差范围越大, 因而过程能力就越低。因此, 可以从过程能力指数的数值大小来判断能力的高低。表 1 给出了过程能力指数与过程能力之间关系。从经济和质量两方面的要求来看, 过程能力指数值并非越大越好, 而应在一个适当的范围内取值。在有过程能力的条件下, 过程能力指数 C_p 的置信水平 $1 - \alpha$ 的贝叶斯单侧置信区间为 $[\hat{C}_{pB}^{(1)}, \infty)$ 。

3 贝叶斯动态评价模型

如果根据第一阶段的信息, 生产过程能力充足, 那么此阶段的样本数据信息可以保存下来, 用于下

一个阶段的过程能力评价分析。假设过了一段时间以后,需要进一步分析生产过程能力,随机抽出 m_2 个样本:

$$x_i^{(2)} = \{ x_{i1}^{(2)}, x_{i2}^{(2)}, \dots, x_{in_i^{(2)}}^{(2)} \}, i = 1, 2, \dots, m_2 \quad (25)$$

表 1 贝叶斯过程能力指数与过程能力的评价

$C_{pb}^{(k)}$ 的取值范围	过程能力的评价	应采取的措施
$C_{pb}^{(k)} > 1.67$	过程能力过剩	(1) 修订标准,缩小公差,保证更高的质量水平;(2) 降低对原材料或机械设备的要 求,放宽检查,设法降低成本。
$1.33 < C_{pb}^{(k)} \leq 1.67$	过程能力充足	(1) 采用控制图控制工序,使其处于稳定状态,并保持过程能力不变;(2) 适当简化产品检 验工作,争取节约管理成本。
$1.00 < C_{pb}^{(k)} \leq 1.33$	有过程能力	(1) 利用控制图来监视工序状态的变化,确保产品质量稳定;(2) 调查机械能力,确认机械 能力充足后,再过渡到工序管理。
$0.67 < C_{pb}^{(k)} \leq 1.00$	过程能力不足	(1) 分层调查影响过程能力的各种因素,并采取相应措施,使其恢复正常;(2) 对产品进行 全数检查,进行分级筛选,确保出厂成品的质量。
$C_{pb}^{(k)} \leq 0.67$	无过程能力	(1) 不能继续生产,必须改革工艺,待确认过程能力充足后,再进行正常生产;(2) 对已生产 的产品进行全数检查。

此处 $n_i^{(2)}$ 为第二阶段第 i 个样本的样本大小;并记

$$X^{(2)} = \{ x_1^{(2)}, x_2^{(2)}, \dots, x_{m_2}^{(2)} \}, m_2 = \sum_{i=1}^{m_2} n_i^{(2)} \quad (26)$$

对于给定的样本 $X^{(2)}$, 参数 μ 和 σ^2 的似然函数为

$$L_2(\mu, \sigma^2 / X^{(2)}) = \prod_{i=1}^{m_2} \prod_{j=1}^{n_i^{(2)}} f(x_{ij}^{(2)} / \mu, \sigma^2) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{n_2/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{m_2} \sum_{j=1}^{n_i^{(2)}} (x_{ij}^{(2)} - \mu)^2\right\} \quad (27)$$

此时,我们选择初试阶段中参数 μ 和 σ^2 的联合后验分布 $\pi_1(\mu, \sigma^2 / X^{(1)})$ 作为第二阶段参数的先验分布,即

$$\pi_2(\mu, \sigma^2) = \pi_1(\mu, \sigma^2 / X^{(1)}), -\infty < \mu < \infty, \sigma^2 > 0 \quad (28)$$

该先验分布的选择正体现了贝叶斯理论的核心

$$\pi_2 = \frac{1}{Z} \prod_{k=1}^2 \prod_{i=1}^{m_k} \prod_{j=1}^{n_i^{(k)}} \frac{1}{\sigma_k \sqrt{2\pi}} \exp\left\{-\frac{(x_{ij}^{(k)} - \bar{x}^{(k)})^2}{2\sigma_k^2}\right\}, \bar{x}^{(2)} = \frac{1}{m_2} \sum_{i=1}^{m_2} \sum_{j=1}^{n_i^{(2)}} x_{ij}^{(2)} \quad (31)$$

同样地,可以证明:精度参数 σ^2 的后验边缘分布仍为伽玛分布,其形状参数为 $\nu_2 = (m_1 + m_2 - 1)/2$, 位置参数为 λ_2 , 即

$$\sigma^2 / X^{(2)} \sim \Gamma(\nu_2, \lambda_2) \quad (32)$$

据此计算 σ^2 的后验期望

$$E(\sigma^2 / X^{(2)}) = \frac{(\nu_2 + 0.5)}{6 \cdot \frac{1}{2} \Gamma(\nu_2)} \quad (33)$$

以及过程能力指数的后验期望和后验方差

$$E(C_p / X^{(2)}) = \frac{(\nu_2 + 0.5)(USL - LSL)}{6 \cdot \frac{1}{2} \Gamma(\nu_2)} \quad (34)$$

$$Var(C_p / X^{(2)})$$

思想:历史数据信息的充分利用;在此先验分布下, μ 和 σ^2 在第二阶段的联合后验分布分布密度函数为

$$\pi_2(\mu, \sigma^2 / X^{(2)}) = \left(\frac{1}{\sigma^2}\right)^{(n_1+n_2+2)/2} \exp\left\{-\frac{1}{2\sigma^2} \left[\sum_{i=1}^{m_1} \sum_{j=1}^{n_i^{(1)}} (x_{ij}^{(1)} - \mu)^2 + \sum_{i=1}^{m_2} \sum_{j=1}^{n_i^{(2)}} (x_{ij}^{(2)} - \mu)^2 \right]\right\} \quad (29)$$

与 2.1 中参数后验边缘分布密度函数的计算方法类似,根据 μ 和 σ^2 的联合后验分布 $\pi_2(\mu, \sigma^2)$, 我们可以推导出方差 σ^2 的后验边缘分布密度函数

$$\pi_2(\sigma^2 / X^{(2)}) = \left(\frac{1}{\sigma^2}\right)^{(n_1+n_2+1)/2} \exp\left(-\frac{1}{\sigma^2}\right) \quad (30)$$

其中

$$\pi_2 = \frac{1}{Z} \prod_{k=1}^2 \prod_{i=1}^{m_k} \prod_{j=1}^{n_i^{(k)}} \frac{1}{\sigma_k \sqrt{2\pi}} \exp\left\{-\frac{(x_{ij}^{(k)} - \bar{x}^{(k)})^2}{2\sigma_k^2}\right\}, \bar{x}^{(2)} = \frac{1}{m_2} \sum_{i=1}^{m_2} \sum_{j=1}^{n_i^{(2)}} x_{ij}^{(2)} \quad (31)$$

$$= \frac{(USL - LSL)^2}{36 \cdot \frac{1}{2} \Gamma(\nu_2)} \left[\frac{1}{\sigma^2} - \left(\frac{\nu_2 + 0.5}{\frac{1}{2} \Gamma(\nu_2)}\right)^2 \right] \quad (35)$$

在二次损失函数下,第二阶段过程能力指数的贝叶斯点估计为其后期期望,即

$$C_{pb} = \frac{(\nu_2 + 0.5)(USL - LSL)}{6 \cdot \frac{1}{2} \Gamma(\nu_2)} \quad (36)$$

另一方面,若记 $G(\nu_2, \lambda_2)$ 为伽玛分布 $\Gamma(\nu_2, \lambda_2)$ 的 α 分位数,则

$$P\{C_p > \frac{USL - LSL}{6} \sqrt{G(\nu_2, \lambda_2)} / X^{(2)}\} = 1 - \alpha \quad (37)$$

因此,在第二阶段中过程能力指数下限的贝叶

斯估计为

$$c_{pB}^{(2)} = \frac{USL - LSL}{6} \sqrt{G(\alpha, \beta)} \quad (38)$$

相应地,置信水平 $1 - \alpha$ 的动态单侧置信区间估计为 $[\frac{USL - LSL}{6} \sqrt{G(\alpha, \beta)}, \infty)$ 。

过程能力的评价步骤与 2.2 相同,并且,根据过程能力指数的取值大小,如果此阶段过程能力充足或过剩,则样本数据信息可以保存下来,作为历史数据信息用于下一阶段过程能力的分析研究。

4 实证分析

作为应用,我们利用上述方法研究某电子企业彩色 STN-LCD 显示屏制造过程能力问题;在此问题中,所要需要考察的产品质量指标为显示屏的厚度(单位:mm),不妨将其设为 X ,其上下限 USL, LSL 分别为 7.50mm 和 6.50mm,目标值为 7.00mm。在第一阶段随机抽出 20 批产品,每批次的样本大相同,均为 10;表 1 为样本数据的具体观察值,表 2 列出了相应的子样均值和标准差。

表 2 第一阶段的样本数据

i	观察值									
1	7.06	6.88	7.05	6.94	7.11	6.90	7.09	6.92	6.89	7.11
2	6.93	6.98	6.99	6.85	6.97	7.01	6.92	6.88	6.92	7.25
3	6.74	7.07	7.02	7.02	6.93	7.20	7.06	7.13	7.21	7.04
4	6.88	6.99	7.13	6.92	7.00	6.87	6.97	7.04	6.90	6.97
5	6.92	7.01	6.92	7.01	7.06	6.97	7.01	6.78	6.89	6.97
6	6.86	6.99	7.02	6.86	7.23	6.97	7.14	6.94	7.17	7.04
7	6.78	6.92	6.94	6.99	7.03	6.99	7.03	6.94	6.94	6.94
8	7.06	7.16	6.96	7.11	6.83	6.98	6.92	6.99	7.14	7.02
9	7.05	6.91	6.91	7.04	7.13	6.92	7.06	6.98	7.00	7.12
10	6.96	7.06	7.07	6.97	6.94	6.92	6.80	6.89	6.88	6.86
11	7.13	6.97	6.74	6.81	7.02	7.03	7.07	7.14	6.87	6.92
12	6.93	7.05	7.10	6.88	6.99	7.01	7.08	7.04	7.18	7.14
13	6.76	7.09	6.81	6.95	6.87	6.93	7.05	6.99	6.93	6.89
14	7.11	7.09	7.13	6.96	6.97	7.24	7.03	6.87	6.86	6.81
15	7.02	7.13	7.20	6.98	7.00	7.12	6.99	6.96	6.95	7.07
16	6.93	6.87	7.03	6.98	7.13	7.00	6.93	7.11	6.95	6.88
17	7.04	7.05	6.98	6.87	7.19	6.78	6.98	7.05	6.97	7.07
18	6.94	7.00	7.01	6.97	7.05	7.05	7.11	6.99	7.17	7.20
19	7.10	6.98	6.86	6.88	7.03	6.81	6.73	6.91	7.03	6.93
20	6.98	7.25	6.97	6.89	6.84	7.16	6.92	7.08	6.97	7.02

表 3 第一阶段样本的子样平均值与标准差

i	\bar{x}_i	s_i	i	\bar{x}_i	s_i
1	6.9950	0.0969	11	6.9700	0.1340
2	6.9700	0.1103	12	7.0400	0.0918
3	7.0420	0.1361	13	6.9270	0.1013
4	6.9670	0.0797	14	7.0070	0.1375
5	6.9540	0.0801	15	7.0420	0.0840
6	7.0220	0.1257	16	6.9810	0.0883
7	6.9500	0.0716	17	6.9980	0.1128
8	7.0170	0.1031	18	7.0490	0.0860
9	7.0120	0.0820	19	6.9260	0.1122
10	6.9350	0.0848	20	7.0080	0.1249

在进行过程能力评价之前,我们先利用样本数据建立均值 - 标准差控制图。图 1 显示:均值控制图和标准差控制图中所有点均处于控制限内,并且

这些点的排列具有随机模式的特点,说明制造过程处于统计控制状态,因此,可以进一步研究过程能力大小问题。

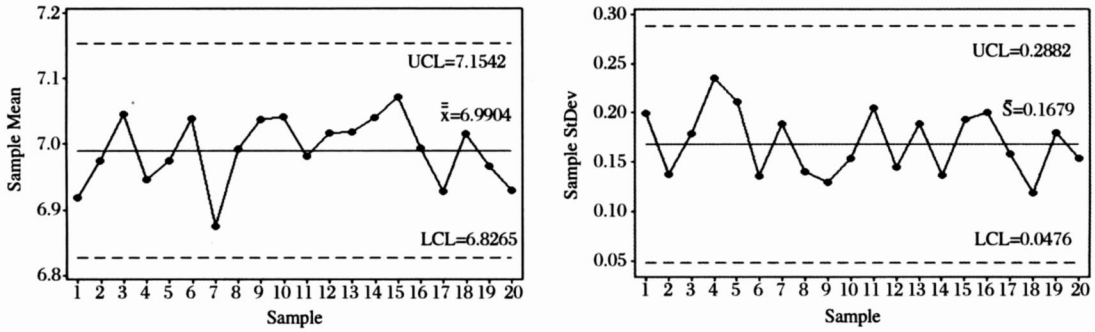


图 1 第一阶段的样本均值 - 标准差控制图

在本例中,形状参数 $\alpha = 99.50$,尺度参数 $\beta = 0.8879$,伽玛分布 $(99.50, 0.8879)$ 的 0.05 分位数为 74.2999;根据(19)和(24)计算 $\hat{C}_{pb}^{(1)}$ 与 $\hat{C}_{pc}^{(1)}$,过程能力指数及其下限的贝叶斯估计分别为 1.5646 和 1.4366。由于 $\hat{C}_{pb}^{(1)} > 1.33$,我们据此认为:生产过程能力充足,其置信水平 0.95 的区间估计为

$[1.4366, \quad)$;并且,样本数据信息可以用于下一阶段的过程能力分析。

在第二阶段,随机抽出 16 批产品,每批次样本大小均为 10;表 4 和表 5 列出了样本观察值和相关统计量的取值。

表 4 第二阶段的样本数据

i	观察值									
1	7.02	6.89	7.00	6.94	7.16	7.07	6.96	6.94	7.11	6.99
2	6.99	7.05	6.84	6.72	6.89	7.14	7.17	7.02	7.02	7.10
3	7.07	6.95	6.98	6.88	6.95	6.90	6.84	7.08	7.11	7.09
4	7.02	7.14	7.08	6.97	6.99	6.88	6.94	7.18	7.03	7.06
5	6.93	7.07	7.04	6.98	6.95	7.10	6.92	7.07	7.09	7.10
6	7.18	7.15	6.84	6.98	6.76	6.72	7.01	7.01	7.07	6.93
7	7.09	7.15	7.17	7.01	6.98	6.97	6.97	6.90	7.06	6.82
8	6.87	6.82	6.88	6.98	6.93	7.04	6.92	7.07	7.13	6.90
9	7.15	7.05	6.97	7.02	6.97	7.06	6.78	7.19	7.25	6.96
10	6.94	7.20	6.86	7.05	6.95	6.97	6.99	6.95	7.05	7.00
11	7.09	7.04	6.91	6.95	6.92	6.82	6.83	6.94	6.96	7.18
12	7.03	7.00	6.91	7.02	7.02	6.87	7.07	7.20	7.06	6.96
13	7.07	6.73	7.21	6.91	7.07	6.95	6.91	6.83	7.14	7.06
14	7.03	6.77	6.90	6.90	7.01	6.71	7.01	6.99	7.08	6.98
15	7.17	7.02	6.88	7.08	7.09	6.93	6.98	6.96	6.93	6.85
16	6.87	7.07	6.94	7.21	6.89	7.08	7.11	7.01	6.99	6.92

表 5 第二阶段样本的子样平均值与标准差

i	\bar{x}_i	s_i	i	\bar{x}_i	s_i
1	7.0080	0.0839	9	7.0400	0.1347
2	6.9940	0.1404	10	6.9960	0.0907
3	6.9850	0.0969	11	6.9640	0.1118
4	7.0290	0.0906	12	7.0140	0.0911
5	7.0250	0.0726	13	6.9880	0.1476
6	6.9650	0.1544	14	6.9380	0.1185
7	7.0120	0.1087	15	6.9890	0.1007
8	6.9540	0.0986	16	7.0090	0.1085

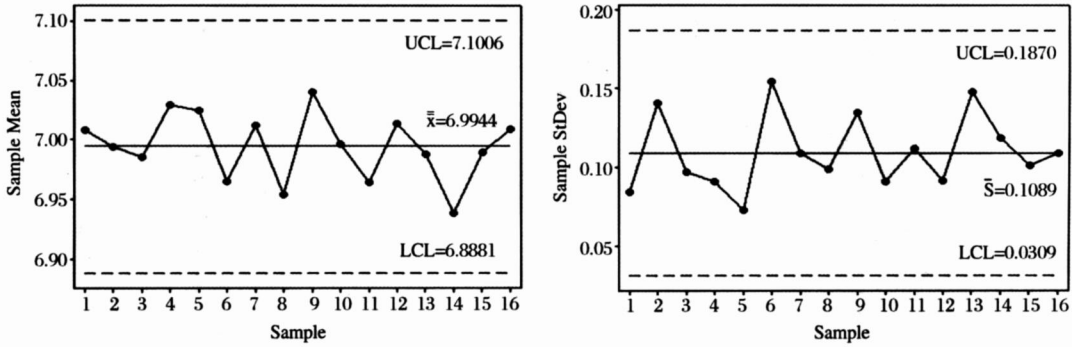


图 2 第二阶段的样本均值 - 标准差控制图

在图 2 中无异常点及异常排列现象,因此生产过程也处于稳定状态,此时相关分布参数如下:

$$\hat{\sigma}^2 = 179.50, \hat{\sigma} = 0.4790, G_{0.05}(179.50, 0.4790) = 75.7037 \quad (39)$$

相应地,其过程能力指数及其下限的贝叶斯估计为:

$$\hat{C}_{pb}^{(2)} = 1.5444, \hat{C}_{pb}^{(2)} = 1.4501 \quad (40)$$

同时,贝叶斯过程能力指数 $\hat{C}_{pb}^{(2)}$ 的后验标准差为 0.00693。由于 $\hat{C}_{pb}^{(2)} > 1.33$,因此第二阶段的生产过程能力充足,其置信水平 0.95 的置信区间为 (1.4501,)。

5 结语

本文通过序贯迭代技术构建了贝叶斯动态过程能力指数估计与评价方法,将过程能力充足条件下的模型参数后验分布作为后一阶段参数的先验分布,据此将历史数据信息转换为参数先验信息;因此,贝叶斯动态过程能力指数评价方法不仅考虑了参数不确定性这一风险影响因素,并且也利用了贝叶斯理论中的信息更新这一核心思想,从而提高了过程能力估计精度。事实上,如果利用基于扩散先验分布的贝叶斯方法进行估计,则第二阶段的过程能力指数的估计值为 1.5138,该估计的标准偏差为 0.0824,而 $\hat{C}_{pb}^{(2)}$ 的标准偏差为 0.0693,前者大于后者,也就是说,贝叶斯动态估计的结果精度要高于前者。

对于其它类型的过程能力指数,例如,例如 C_{pm}, C_{pk} 和多变量过程能力指数,我们也可以探讨其贝叶斯动态估计问题,研究思路与本文无实质性的差异;并且,也可以结合过程质量监控和过程能力分析,构建贝叶斯动态过程质量监控与过程能力的综合分析方法。这些工作都是质量管理领域值得研究的重要问题,也是作者后续的研究工作。

参考文献:

- [1] 胡华安,田益峰,向卓元,邓望年. 物资采购综合评价指标体系质量控制模型应用研究[J]. 中国管理科学, 2004,12(1):75 - 80.
- [2] Juran, J. M. . Quality Control Handbook[M]. McGraw - Hill, New York,1974.
- [3] Chan, L. K. , Cheng, S. W. , Spiring, F. A. . A new measure of process capability: Cpm [J]. Journal of Quality Technology, 1988, 20(2):162 - 175.
- [4] Gupta, A. K. , Kotz, S. . A. . New Process Capability Index[J]. Metrika, 1997, 45(2):213 - 0224.
- [5] Boyle, R. A. . The Taguchi capability index[J]. Journal of Quality Technology, 1991, 23(2):107 - 126.
- [6] Wang, C. H. . Constructing multivariate process capability indices for short-run production [J]. International Journal of Advanced Manufacturing Technology, 2005, 26(21):1306 - 1311.
- [7] 孔祥芬,何桢,宗志宇. 基于矩法的非正态过程能力指数估算方法的研究[J]. 工业工程, 2007(1):57 - 59.
- [8] Chen, X. Y. , Zhu, L. J. . , Shen, X. J. . Accurate calculation of process capability Indices considering the skewed distribution[J]. Journal of Shanghai University, 2001(1):7 - 10.
- [9] 李元生,孙静,祝伟. 一类基于非对称偏差的过程能力指数[J]. 数理统计与管理, 2005, 24(2):35 - 39.
- [10] 余忠华,郑希俊,郑小林. 基于约翰逊曲线拟合的非正态工序能力指数估计[J]. 机械科学与技术, 2002, 21(6):878 - 880.
- [11] 田志友,田澎,王浣尘. 基于 Bootstrap 抽样的多元过程能力指数估计[J]. 管理工程学报, 2006, 20(2):74 - 77.
- [12] Lorenc, A. C. , Hammon, O. Objective quality-control of observations using Bayesian methods - theory, and apractical implementation[J]. Quarterly Journal of the Royal Meteorological Society, 1988, 114(480):513

- 543.
- [13] Bernardo, J. M. , Irony, T. Z. . A general multivariate Bayesian process capability index[J]. The Statistician, 1996, 45(4):487 - 502.
- [14] 林静,韩玉启,朱慧明. 基于 MCMC 稳态模拟的贝叶斯经验费率厘定信用模型[J]. 中国管理科学, 2006, 14(2):33 - 38.
- [15] 朱慧明,赵锐. 基于自相关过程的贝叶斯质量控制模型研究[J]. 计算机集成制造系统, 2008, 14(3):615 - 624.
- [16] Cheng, S. W. , Spiring, F. A. . Assessing process capability - A Bayesian approach[J]. IIE Transactions, 1989, 21(1):97 - 98.
- [17] Shiau, J. J. H. , Chiang, C. T. , Hung, N. A. . Bayesian procedure for process capability assessment [J]. Quality and Reliability Engineering International, 1999, 15(5):369 - 378.
- [18] Shiau. J. J. H. , Hung, H. N. , Chiang, C. T. . A note on Bayesian estimation of process capability indices[J]. Statistics & Probability Letters, 1999, 45(3):215 - 224.
- [19] Wu, C. W. . Assessing process capability based on Bayesian approach with subsamples [J]. European Journal of Operation Research, 2008, 184(1):207 - 228.
- [20] Singh, H. P. , Saxena, S. . Bayesian and shrinkage estimation of process capability index Cp[J]. Communications in Statistics - Theory and Methods, 2005, 34(1):205 - 228.
- [21] Bernardo, J. M. , Irony, T. Z. . A general multivariate Bayesian process capability index[J]. The Statistician, 1996, 45(4):487 - 502.
- [22] 朱慧明,韩玉启. 基于随机参数的贝叶斯统计质量控制模型[J]. 哈尔滨工业大学学报, 2005, 38(4):498 - 501.
- [23] Niverthi. M. , Dey, D. K. . Multivariate process capability a Bayesian perspective [J]. Communications in Statistics - Simulation and Computation, 2000, 29(2):667 - 687.

Bayesian Dynamic Estimation and Evaluation of Process Capability Indices with Multiple Subsamples

ZHU Hui-ming¹, ZENG Hui-fang¹, YU Ke-ming², HAO Li-ya¹, LI Su-fang¹

(1. College of Business Administration, Hunan University, Changsha 410082, China;

2. Department of Statistics, Brunel University, London UB8 3PH, China)

Abstract : To analyze the process capability under random parameters, a new kind of process capability index is designed in this paper. Based on the statistical model for quality variables, we explored the parameters' Bayesian point estimates and interval estimation with a diffuse prior and developed a Bayesian process capability index. Then, we considered the parameters' current posterior distribution to be their prior distribution in the next phrase when the process is capable, by which the Bayesian dynamic process capable index is established. Finally, we give an example to show how to use the method proposed in this paper. The results indicate that the accuracy of estimators for process capability indices can be improved through Bayesian dynamic statistical methods.

Key words : quality control ; process capability index ; bayesian method ; estimation ; prior distribution