

用 Bayes 方法修正质量控制界限*

张连弟 韩效宥

(山西财经大学, 太原030012)

摘 要: 在工序质量控制中, 画控制图的传统作法是用样本均值近似代替总体的均值. 由于样本的随机性, 因此, 估计误差在所难免. 本文用 Bayes 方法将用样本均值 $\bar{X}$ 代替总体均值 μ 所产生的估计误差一并计入到总体 X 所固有的随机性之中, 从而修正了传统的质量控制界限的画法.

关键词: 质量控制, 控制界限, Bayes 方法

分类号: AM S(1991) 62N 10/CL C O 212 1

文献标识码: A **文章编号:** 1000-341X(1999)增刊-0312-03

在概率论中知道, 当生产工序不存在系统误差时, 产品的质量特性值(总体 X)服从正态分布 $N(\mu, \sigma^2)$. 那么, 样品的质量特性值 X 落在区间 $(\mu - 3\sigma, \mu + 3\sigma)$ 中的概率为 0.997. 据此(所谓的“ 3σ 准则”), 画出工序质量控制图的中心线(CL_X)及上, 下控制界限(UCL_X, LCL_X).

$$\begin{cases} CL_X = \mu, \\ UCL_X = \mu + 3\sigma, \\ LCL_X = \mu - 3\sigma \end{cases} \quad (1)$$

在实际操作中, 总体分布中的方差 σ^2 往往是已知的, 而 μ 未知. 为了画出控制界限, 用稳定状态下的部分产品(简单随机抽样)的质量特性值 $X_1, X_2, \dots, X_n$, 求得的样本均值 $\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n)$ 近似代替 μ 来确定 CL_X, UCL_X 及 LCL_X 的近似位置:

$$\begin{cases} CL_X \approx \bar{X}, \\ UCL_X \approx \bar{X} + 3\sigma, \\ LCL_X \approx \bar{X} - 3\sigma \end{cases} \quad (2)$$

但是, 由于样本 $X_1, X_2, \dots, X_n$ 的随机性, 因此, 用 $\bar{X}$ 估计 μ 而产生的估计误差在所难免. 于是, 产品质量特性值 X 落在 $(\bar{X} - 3\sigma, \bar{X} + 3\sigma)$ 内的概率一般不等于 0.997. 为了确保用 $\bar{X}$ 画控制界限的准确程度, 使画出的控制图更加合理, 用 Bayes 方法对 (2) 式进行修正.

古典统计观点是假定总体分布中待估参数 μ 是未知的常数, 而 Bayes 方法则将待估参数 μ 视为随机变量. 人们可以根据过去的资料(或经验)人为确定参数 μ 的概率密度 $\varphi(\mu)$, 将其称为 μ 的先验分布. 设 $X_1, X_2, \dots, X_n$ 为从总体 X 中抽得的一个样本, 记该抽样结果为 ϵ . 那么, μ 的先验分布密度 $\varphi(\mu)$ 可以根据 ϵ 提供的信息利用 Bayes 方法予以修正而求得后验分布密度

* 收稿日期: 1996-10-03

$$\mathcal{Q}(\mu). \text{ 由 } \mathcal{Q}(\mu_i) \Delta\mu \approx \frac{P(\epsilon|\mu_i < \mu \leq \mu_i + \Delta\mu) \cdot \mathcal{Q}(\mu_i) \Delta\mu}{\sum_k P(\epsilon|\mu_k < \mu \leq \mu_k + \Delta\mu) \cdot \mathcal{Q}(\mu_k) \Delta\mu}, \text{ 当 } \Delta\mu \rightarrow 0 \text{ 时, 上式取极限得}$$

$$\mathcal{Q}(\mu) = \frac{P(\epsilon|\mu) \cdot \mathcal{Q}(\mu)}{\int_{-\infty}^{\infty} P(\epsilon|\mu) \cdot \mathcal{Q}(\mu) d\mu}. \quad (3)$$

由于总体 $X \sim N(\mu, \sigma^2)$. 故

$$P(\epsilon|\mu) = \prod_{k=1}^n \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\chi_k - \mu)^2}{2\sigma^2}} \cdot d\chi = \prod_{k=1}^n N_{\mu}(\chi_k, \sigma^2) d\chi, \quad (4)$$

符号 $N_{\mu}(\chi_k, \sigma^2)$ 表示均值为 χ_k , 方差为 σ^2 的正态随机变量 μ 的分布密度 于是

$$\begin{aligned} \mathcal{Q}(\mu) &= \frac{[\prod_{k=1}^n N_{\mu}(\chi_k, \sigma^2)] \cdot \mathcal{Q}(\mu)}{\int_{-\infty}^{\infty} [\prod_{k=1}^n N_{\mu}(\chi_k, \sigma^2)] \cdot \mathcal{Q}(\mu) d\mu} \\ &= \frac{N_{\mu}(\bar{X}, \sigma^2/n) \cdot \mathcal{Q}(\mu)}{\int_{-\infty}^{\infty} N_{\mu}(\bar{X}, \sigma^2/n) \cdot \mathcal{Q}(\mu) d\mu} = m \cdot N_{\mu}(\bar{X}, \sigma^2/n) \cdot \mathcal{Q}(\mu), \end{aligned} \quad (5)$$

其中 $m = [\int_{-\infty}^{\infty} N_{\mu}(\bar{X}, \sigma^2/n) \cdot \mathcal{Q}(\mu) d\mu]^{-1}$ 为常数; $\bar{X} = \frac{1}{n}(\chi_1 + \chi_2 + \cdots + \chi_n)$.

有了由观测数据 $\chi_1, \chi_2, \cdots, \chi_n$ 得到的关于参数 μ 的后验分布密度 $\mathcal{Q}(\mu)$, 就可以将用 $\bar{X}$ 估计 μ 所产生的误差与随机变量 X 固有的随机性结合起来, 利用全概率公式一并计入到随机变量 X 之中. 此时, 随机变量 X 的分布密度为

$$f_X(X) = \int_{-\infty}^{\infty} f_X(X|\mu) \mathcal{Q}(\mu) d\mu \quad (6)$$

事实上, 设 $F(X)$ 为随机变量 X 的分布函数, 由于

$$F(X) = P(X \leq X) \approx \sum_k P(X \leq X|\mu_k < \mu \leq \mu_k + \Delta\mu) \cdot \mathcal{Q}(\mu_k) \Delta\mu,$$

于是, 当 $\Delta\mu \rightarrow 0$ 时, 得

$$F(X) = \int_{-\infty}^{\infty} P(X \leq X|\mu) \cdot \mathcal{Q}(\mu) d\mu$$

将上式对 X 求导即得(6)式

下面根据实际中经常遇到的两种情况, 具体讨论 $f_X(X)$ 的表达式(6), 进而给出用 $\bar{X}$ 代替 μ 时(2)式的修正结果

(一) 若没有关于 μ 的资料(或信息)提供时, 则可以认为 μ 的取值是均匀分布的 为此, 可假定 $\mathcal{Q}(\mu) = 1$, 于是, 由(5)式得 $\mathcal{Q}(\mu) = m \cdot N_{\mu}(\bar{X}, \sigma^2/n)$. 既然 $\mathcal{Q}(\mu)$ 是随机变量的分布密度, 归一化后, m 必然等于1. 因此,

$$\mathcal{Q}(\mu) = N_{\mu}(\bar{X}, \sigma^2/n). \quad (7)$$

将(7)式代入(6)式, 得

$$f_X(X) = \int_{-\infty}^{\infty} N_X(\mu, \sigma^2) \cdot N_{\mu}(\bar{X}, \sigma^2/n) d\mu = N_X(\bar{X}, \sigma^2 + \sigma^2/n).$$

这就是说 $X \sim N(\bar{X}, \sigma^2 + \sigma^2/n)$. 因此, 根据“ 3σ 准则”写出 X 的控制界限为

$$\begin{cases} \text{CL}_X = \bar{X} \\ \text{UCL}_X = \bar{X} + 3\sqrt{\sigma^2 + \sigma^2/n} \\ \text{LCL}_X = \bar{X} - 3\sqrt{\sigma^2 + \sigma^2/n} \end{cases} \quad (8)$$

由此可见,对于(一)这种数学模型,与(2)式比较,工序控制中心线 CL_X 没变,而上、下控制界限 $\text{UCL}_X, \text{LCL}_X$ 的位置发生变化

(二) 当 $\mu \sim N(\mu_1, \sigma_1^2)$ 时, (其中 μ_1 及 σ_1^2 均为已知常数), 则

$$\mathcal{Q}_2(\mu) = m \cdot N_\mu(\bar{X}, \sigma^2/n) \cdot N_\mu(\mu_1, \sigma_1^2) = N_\mu(\mu_2, \sigma_2^2).$$

其中 $\mu_2 = \frac{\bar{X} \cdot \sigma_1^2 + \mu_1 \cdot \sigma^2/n}{\sigma_1^2 + \sigma^2/n}$, $\sigma_2^2 = \frac{\sigma_1^2 \cdot \sigma^2/n}{\sigma_1^2 + \sigma^2/n}$. 将上述 $\mathcal{Q}_2(\mu)$ 代入(6)式, 得:

$$f_X(X) = \int_{-\infty}^{\infty} N_X(\mu, \sigma^2) \cdot N_\mu(\mu_2, \sigma_2^2) d\mu = N_X(\mu_2, \sigma^2 + \sigma_2^2),$$

即: $X \sim N(\mu_2, \sigma^2 + \sigma_2^2)$, 因此, X 的控制界限为

$$\begin{cases} \text{CL}_X = \mu_2, \\ \text{UCL}_X = \mu_2 + 3\sqrt{\sigma^2 + \sigma_2^2}, \\ \text{LCL}_X = \mu_2 - 3\sqrt{\sigma^2 + \sigma_2^2} \end{cases} \quad (9)$$

从这个结果看, 与(2)式比较, 对于(二)这种数学模型, $\text{CL}_X, \text{UCL}_X$ 及 LCL_X 的位置均发生了变化

参 考 文 献

- [1] 冯士雍等编 质量管理统计方法 [M] 北京: 机械工业出版社, 1984
- [2] 中科院数学研究所统计组编 抽样检验方法 [M] 北京: 科学出版社, 1977.
- [3] 黄克中, 毛善培编著 随机方法与模糊数学应用 [M] 上海: 同济大学出版社, 1987.
- [4] 吴乙申主编 应用统计学 [M] 北京: 机械工业出版社, 1986

An Application of Bayesian Method to Correct Quality Control Limits

Zhang L ian di H an X iaoyou

(Shanxi University of Finance and Economics, Taiyuan 030012)

Abstract

Traditional way of plotting control chart in the process quality control is the use of sample mean instead of population mean. Estimating error is hard to avoid because of the sample's randomness. In this paper, we use Bayesian method to consider estimating error, which is produced by sample mean $\bar{X}$ replacing population mean μ , belonging to fixed randomness of population rv X , thus correcting the traditional way of quality control limits.

Keywords quality control, control limits, Bayesian method