

研发项目工作量估计方法的可靠性评估与预测

张俊光, 杨芳芳, 徐振超

(北京科技大学 东凌经济管理学院, 北京 100083)

摘要:为提高研发项目工作量估计的精确度,减少因工作量估计方法应用不当造成的研发项目失败,将一种基于新 Dirichlet 先验分布的适合小子样复杂系统可靠性增长评估与预测的 Bayesian 模型应用于验证研发项目工作量估计方法的有效性。通过利用先验信息和阶段试验信息,为先验参数的确定提供了定量方法;利用 Gibbs 抽样合理估算出后续阶段的有效性,实现了未来阶段工作量估计方法有效性的预测,为研发企业采纳工作量估计方法提供了决策依据。

关键词:研发项目;工作量估计;Bayesian 模型;产品可靠性验证;Gibbs 抽样

DOI:10.6049/kjbydc.2012040243

中图分类号:G311

文献标识码:A

文章编号:1001-7348(2013)06-0118-03

0 引言

为提高产品可靠性,在研制新产品期间,通常会经历试验并改进的反复过程,即可靠性验证试验。近年来,已有相当多的学者对如何进行可靠性验证试验进行了研究分析,以评估各种产品的可靠性。

郭海涛等^[1]运用 Markov 模型对安全系统的可靠性进行了定量评估;Yuanming Huang^[2]则基于 WSEI-AC 模型评估了安全系统的有效性;马溪原等^[3]基于故障影响遍历算法的可靠性评估方法,通过建立区域网络图、节点邻接矩阵、故障影响矩阵,对含微电网的复杂配电网进行了可靠性评估;斗计华等^[4]运用模糊综合评价方法,评定武器系统单个设备的使用可靠度;何明等^[5]提出一种改进二元决策图的网络可靠性评估方法。现有大多数研究都是对产品和设备的可靠性进行检测,只有少数学者将可靠性验证运用于检测某种新的技术和方法的可靠性上。例如,常士楠等^[6]用数值仿真方法对 AEDC 结冰缩比方法的有效性进行了验证。但是,目前未见关于管理方法可靠性检验的研究,更没有验证研发项目工作量估计方法有效性的相关研究。

1 研发项目工作量估计方法有效性验证的必要性

据 Standish Group 2004 年对 50 000 多个研发项目的统计结果显示,项目平均预算超支 89%,对研发成本估算不足是其主要原因之一^[7]。对研发项目工作量估算不准确,已经成为研发项目失败的重要原因^[8]。因此,寻求一种有效性较高的工作量估算方法是各研发企业迫切希望解决的问题。

随着研发行业的发展与科学研究的深化,各种各样的工作量估计方法被提出,包括参数模型法、模糊逻辑法^[9]、神经网络模型^[10]、COCOMO 模型、组合预测方法^[11]等。一方面,研发企业期望运用新方法来改变目前面临的工作量估计不准确问题;另一方面,每一种新的估算方法都是有其风险的,现有研究提出的诸多工作量估计方法,没有一种是适应于所有企业的,且运用不同方法得出的结果之间的差异十分显著^[12]。一种有效性低的工作量估计方法的运用和推广必然导致项目的失败。同时,由于研发项目一般具有投资大、周期长等特点,这就意味着研发项目的失败将给研发企业带来极大损失。因此,研发公司在大范围推广某种研发项目工作量估计方法之前,对方法的有效性进行预测

收稿日期:2012-06-11

基金项目:教育部人文社科基金项目(10YJC630384);中央高校基本科研业务费项目(FRF-BR-12-007,020);国家自然科学基金项目(70972120)

作者简介:张俊光(1972-),男,山东即墨人,博士,北京科技大学东凌经济管理学院副教授,研究方向为项目管理;杨芳芳(1988-),女,贵州贵平人,北京科技大学东凌经济管理学院硕士研究生,研究方向为软件项目管理;徐振超(1988-),女,山东聊城人,北京科技大学东凌经济管理学院硕士研究生,研究方向为软件项目风险管理。

与研究是很有必要的。

然而, 研发项目不同于一般产品, 在一种工作量估计方法尚未被证实具有较高有效性之前是不可能很多项目中同时试验的, 这就决定了工作量估计方法有效性测试的小样本性。另外, 对工作量估计方法有效性进行研究的最重要目的是预测下一阶段该方法的可靠性, 以此作为是否运用与推广该方法的决策依据。

现有大部分可靠性验证模型一般仅能给出当前阶段试验产品的可靠性估计, 无法预测最终产品的可靠性。为解决这一问题, 明志茂等^[13]基于新 Dirichlet 先验分布, 建立了一种适合小子样复杂系统异总体可靠性增长分析的 Bayesian 模型。该模型能够合理估算出当前阶段和后续试验阶段的产品可靠性, 可以实现对未来阶段可靠性的预测。因此, 将该模型运用于工作量估算方法的有效性测试, 能够为企业决策提供理论依据与数据支持。

2 工作量估算方法的有效性评估与预测模型

假设某个工作量估计方法在正式大范围推广于企业的研发项目工作量估算之前, 会经过 n 个试验阶段, 每个试验阶段将该方法运用于 m_k 个项目中, 其中有 s_k 个项目成功, 则可以得到第 n 个阶段的方法有效性数据, 并可计算出该阶段的方法有效性 R_k 。与用于检验产品生产的可靠性试验分析数据不同的是, 在研发项目工作量估计方法的试验中, 每个阶段用于测试的项目个数, 即 m_k 的数量很小, 这是由项目的大型性与高投入等特殊性质决定的。另外, 在对方法进行试验的试验中, 一种工作量估计方法成功与否的定义也与产品不相同。对于研发项目工作量估计方法来说, 由于目前普遍的估计方法其误差较大, 项目公司可以依据自身对项目的要求, 或依据研发企业的过程能力基准 (PCB), 来判定项目实际工作量与估计工作量之间的偏差达到多大范围才算项目成功。

具体的试验操作过程如下:

第 1 阶段, 运用工作量估计方法对 m_1 个项目估算其各自的工作量, 在项目结束之后, 对比实际工作量与计划工作量之间的偏差, 偏差在可接受的范围内则认为项目成功, 据此统计得出成功项目的个数 s_1 , 计算出本阶段的可靠性 R_1 。同时, 总结项目经验, 找出项目失败的原因并对估计方法进行改进。

第 2 阶段, 运用该工作量估计方法对 m_2 个项目估算其各自的工作量, 在项目结束之后, 统计得出成功项目的个数 s_2 , 计算得出本阶段的可靠 R_2 。一般来说, 经过改进后的方法其可靠性会在一定程度上比原方法高, 即 $R_2 > R_1$, 若此时得出的项目可靠性达到公司要求, 即可停止试验, 否则再次对估计方法进行改进。

第 n 阶段, 运用相同的方法计算出本阶段的可靠性 R_n , 且有 $R_n > \dots > R_k > \dots > R_2 > R_1$ 。此时, 经过软件

分析与预测可知 R_{n+1} 已经达到研发公司可接受的可靠性范围, 因此可以停止试验, 并在下一阶段将此工作量估计方法大量推广于公司的工作量估算中。

假定某公司规定项目的实际工作量与计划工作量之间的偏差在 10% 以内, 即认为项目的工作量估计方法是成功的。另外, 该企业规定, 一个有效性达到 0.9 以上的工作量估算方法就是相对成熟的方法, 值得在其它研发项目中大规模推广与使用。

现有一种新的工作量估计方法, 该方法在理论上具有可行性与先进性, 为了检验该方法的有效性, 该公司将对此方法进行有效性试验, 在预测出方法的可靠性达到标准水平 0.9 以上时, 才可在企业内推广。

为了能够估算出未来阶段的可靠性, 通常需要由专家经验或类似项目信息得出新估计方法的先验估计。假设根据以往经验, 一种工作量估算方法要达到有效性 0.9 的目标必须经过 4 个试验阶段, 如表 1 所示, 专家给出各阶段的估计方法可靠性的先验估计 $R = (0.5, 0.7, 0.8, 0.9)$ 。同时, 为了更精确地描述验前信息, 可以给出各试验阶段可靠度的区间估计。其中, 由于在第 1 阶段还没有关于方法的试验和改进措施的信息, 对方法的可靠度把握不大, 专家给出的估计区间长度较大。在进行试验和采取相应的改进措施之后, 专家给出的区间估计变得更为准确, 区间长度逐渐减小。如果目前做了 3 个阶段的增长试验, 各阶段试验结果 (m_k, s_k) 分别为 $(3, 1), (4, 3), (5, 4)$, 利用模型可以对前 3 个阶段的可靠性进行评估, 同时预测第 4 阶段的可靠性。

表 1 试验信息和验前信息

阶段 k	试验信息		验前信息	
	m_k	s_k	R_k	(R_k, L, R_k, H)
1	3	1	0.5	(0.35, 0.65)
2	4	3	0.7	(0.6, 0.8)
3	5	4	0.8	(0.72, 0.88)
4	—	—	0.9	(0.87, 0.93)

资料来源: 普天集团某下属软件公司

基于新 Dirichlet 先验分布的可靠性增长模型假设, 在第 k 检测区间内进行定时截尾试验得到的成败型数据, 其可靠度 R_k 服从参数为 (m_k, s_k) 的二项分布, 似然函数为:

$$L(R_k; m_k; s_k) = \binom{m_k}{s_k} R_k^{s_k} (1 - R_k)^{m_k - s_k} \quad (1)$$

模型通过构造 $(R_{k-1}, 1)$ 上的截尾 Beta 分布来检测第 k 阶段的可靠度。假设第 k 个试验阶段的先验参数 $a_k > 0, b_k > 0$, 则 $B(a_k, b_k)$ 为 Beta 函数。那么, $\tilde{R} = (R_1, R_2, \dots, R_{n-1}, R_n)$ 的联合先验密度函数为

$$\pi(\tilde{R} | \tilde{\alpha}, \tilde{\beta}) \prod_{k=1}^{n+1} g_k(R_k | R_{k-1}) \quad (2)$$

在给定的 R_1, R_2, \dots, R_{k-1} 的条件下, 可以求得 R_k 的条件均值和条件方差为

$$u_k = \frac{a_k + b_k R_{k-1} - 1}{a_k + b_k} v_k = \frac{a_k b_k (1 - R_{k-1})^2}{(a_k + b_k)^2 (a_k + b_k)} \quad (3)$$

为了计算出先验参数的值,可以依据验前信息给出的可靠性区间构建均匀分布,然后以先验参数为变量,将均值作为约束,方差作为目标,利用最优化方法求出与该均匀分布最为接近的 Beta 分布,作为方法验证的先验分布,如图 1 所示。

则均匀分布的均值和方差为

$$ER_k = \frac{R_{k,L} + R_{k,H}}{2} \quad (4)$$

$$VR_k = \frac{(R_{k,H} - R_{k,L})^2}{12} \quad (5)$$

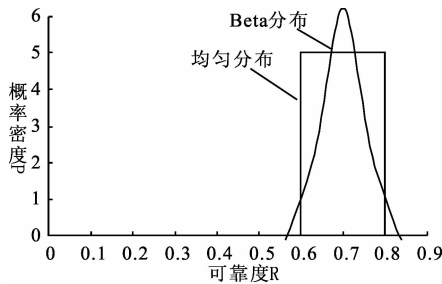


图 1 均匀分布与一般 Beta 分布等效图

通过研究,结合式(3)、式(4)和式(5),采用式(6)的最优化模型可求解第 k 个试验阶段的先验参数

$$\begin{aligned} \min(v_k - VR_k)^2 \\ s. t. u_k = ER_k \quad a_k > 0 \quad b_k > 0 \end{aligned} \quad (6)$$

基于表 1 中的数据,利用式(4)表示的最优化模型可得到各试验阶段与均匀分布 $U(R_{k,L}, R_{k,H})$ 等价的 Beta 分布的先验参数 a_k 和 b_k ,如表 2 所示。另外,表 2 还显示了均匀分布和相应的等价 Beta 分布的均值及方差。

表 2 均匀分布与等效 Beta 分布参数

阶段	均匀分布		等效 Beta 分布			
	ER_k	V_k	ER_k	V_k	a_k	b_k
1	0.5	0.007 5	0.5	0.007 5	16.166 7	16.166 7
2	0.7	0.003 3	0.7	0.003 1	6.872 7	10.309 1
3	0.8	0.002 1	0.8	0.002 1	2.841 3	5.682 5
4	0.9	0.000 3	0.9	0.000 3	16.166 7	16.166 7

在得到上述先验参数后,模型利用 Gibbs 抽样计算可靠度值。Gibbs 抽样算法是一种广泛应用的 Metropolis-Hastings 抽样方法,它通过建立 Markov 链,对未知变量进行模拟,当链达到平稳时,收敛到目标分布,即得到所求的后验分布^[14]。在 Gibbs 抽样中取 10 000 次抽样迭代,可得到各阶段该方法有效性的后验估计结果,如表 3 所示。

表 3 各试验阶段可靠度预测

	均值	标准差	10%	中位点	90%
r[1]	0.497 1	0.081 28	0.390 2	0.497 3	0.603 0
r[2]	0.705 7	0.070 24	0.613 7	0.709 5	0.792 9
r[3]	0.808 8	0.060 77	0.728 9	0.813 6	0.883 7
r[4]	0.906 0	0.034 21	0.860 5	0.909 3	0.947 0

由表 3 可见,根据专家给出的经验信息和前 3 阶段的试验数据,在没有进行第 4 个试验阶段的情况下,利

用 Bayesian 可靠性增长推断模型能够预测后续阶段的可靠性增长。由表 3 中对第 4 阶段的可靠度预测可知,第 4 阶段可靠度的均值达到 0.906,高于该公司规定的 0.9,因此可认为在第 4 阶段该方法的估算结果与公司要求相符,可以将此种工作量估算方法在正式项目中推广。

3 结语

研发项目的特殊性质,决定了研发项目的工作量估算方法必须具有较高的有效性,才能在项目管理过程中应用与推广,有效性低的工作量估计方法将会直接导致项目的失败,并给研发企业带来重大损失,如何评估并预测工作量估计方法的有效性成为研发企业面临的难题之一。本文基于当前试验数据和先验信息,将一种基于新 Dirichlet 先验分布的可靠性增长模型,应用于预测未来阶段研发项目工作量估算方法的有效性。该方法简单易行,且通过与实际项目数据对比,其误差较小,为研发企业采纳和运用新的研发项目管理方法提供了理论依据。该模型不仅适用于对研发项目工作量估算方法的有效性验证,同样也可用于对其它研发项目管理方法,如工作量控制方法、风险管理方法等的有效性验证。

本文的不足之处在于,并没有将模型后续阶段的实际运行结果纳入,以更加精确地定量评估模型的可靠性。下一步的改进之处在于持续跟踪所验证的工作量估计方法在推广过程中的精确性,不断进行方法改进,以进一步提高估计的精度,并将可靠性评估与预测作为一个动态过程,不仅在试验阶段运用,在方法推广运用阶段,也可以根据获得的更多数据对方法的可靠性进行持续评估与预测。

参考文献:

- [1] 郭海涛,阳宪惠.安全系统定量可靠性评估的 Markov 模型[J].清华大学学报:自然科学版,2008,48(1):39-42.
- [2] YUANMING HUANG. Effectiveness evaluation for security system based on WSEIAC model [C]. 2010 3rd IEEE International Conference on Computer Science and Information Technology, 2010:192-195.
- [3] 马溪原,吴耀文,方华亮,等.基于可靠性评估的微电网配置方法[J].电力系统自动化,2011,35(9):73-78.
- [4] 斗计华,陈万春,钟志通.舰空导弹武器系统使用可靠性评估[J].系统工程与电子技术,2011,33(4):954-957.
- [5] 何明,权冀川,郑翔,等.基于二元决策图的网络可靠性评估[J].控制与决策,2011,26(1):32-36.
- [6] 常士楠,洪海华,张玉珠.基于 Dirichlet 先验分布的 Bayes 二项可靠性增长方法[J].系统工程理论与实践,2006(1):131-135.
- [7] 涂燕,魏法杰.软件开发劳动计量方法与软件成本估算[J].中国管理信息化,2010,13(13):40-43.