

DB37

山东省地方标准

DB37/T 2791—2016

软件测试成本度量基本方法

2016-05-13 发布

2016-06-13 实施

山东省质量技术监督局 发布

目 次

前言 III

引言 IV

1 范围 1

2 规范性引用文件 1

3 术语与定义 1

4 软件测试成本构成 2

5 软件测试成本调整因子 4

6 软件测试成本度量 7

前 言

本标准按照GB/T 1.1—2009给出的规则起草。

本标准由济南市质量技术监督局提出

本标准由山东省信息标准化技术委员会归口。

本标准起草单位：国家网络软件产品质量监督检验中心（济南）、山东大学、浪潮软件集团有限公司软件评测实验室、山东师范大学。

本标准主要起草人：田辉、刘善田、王玲、张伟、李学庆、侯卫峰、张化祥、柳平增、孟宪勇、宋士涛、王岳、张燕、张佳佳、郑希源、李国勇、刘文杰。

引 言

本标准用于规范软件测试涉及的各方在软件测试成本度量方法上达成一致。

本标准不包含软件测试成本度量过程中所需使用的各种基准数据或估算模型,相关各方在使用本标准时,应参考权威部门发布的最新基准数据、估算模型开展软件测试成本度量相关活动。

本标准不涉及软件定价,但相关各方可参照本标准测算软件测试成本,从而为软件测试定价提供依据。

软件测试成本度量基本方法

1 范围

本标准规定了软件测试成本的构成、软件测试成本的调整因子和软件测试成本的度量。

本标准适用于软件生存周期支持过程中的成本度量，包括软件测试项目的成本预算、项目决算以及软件测试相关合同的编制。

本标准的适用对象为开发方的软件测试组以及第三方的软件测试组织。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅所注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 11457 信息技术 软件工程术语

GB/T 15532—2008 计算机软件测试规范

GB/T 16260.1—2006 软件工程 产品质量 第1部分：质量模型

GB/T 18492—2001 信息技术 系统及软件完整性级别

GB/T 25000.51 软件工程 软件产品质量要求和评价（SQuaRE） 商业现货（COTS）软件产品的质量要求和测试细则

SJ/T 11463—2013 软件研发成本度量规范

ISO/IEC 19761 软件工程 COSMIC:一种功能规模度量方法(Software engineering -- COSMIC: a functional size measurement method)

ISO/IEC 20926 软件和系统工程 软件度量 IFPUG功能规模度量方法(Software and systems engineering -- Software measurement -- IFPUG functional size measurement method)

ISO/IEC 20968 软件工程 Mk II功能点分析 计数实践手册(Software engineering -- Mk II Function Point Analysis -- Counting Practices Manual)

ISO/IEC 24570 软件工程 NESMA功能 规模度量方法2.1版 功能点分析应用定义和计数指南(Software engineering -- NESMA functional size measurement method version 2.1 -- Definitions and counting guidelines for the application of Function Point Analysis)

ISO/IEC 29881 信息技术 系统和软件工程 FiSMA1.1功能规模度量方法(Information technology -- Systems and software engineering -- FiSMA 1.1 functional size measurement method)

3 术语与定义

GB/T 11457、GB/T 18492—2001和GB/T 25000.51中界定的以及下列术语和定义适用于本文件。

3.1

软件测试成本 software testing cost

为达成软件测试项目目标所需付出的各种资源代价总和，资源包括人、财、物、信息等。

3.2

自动化测试 automatic testing

自动化测试是机器模拟人为驱动，自动执行测试行为的一种过程。

3.3

直接成本 direct cost

与具体软件测试项目直接相关，为达成项目目标而直接付出的各种资源代价总和。

3.4

间接成本 indirect cost

不能具体到某个项目，但与达成软件测试项目目标相关，可以分摊到各个具体项目中的各种资源代价总和。

3.5

人力成本 human resource cost

为达成软件测试项目目标所需付出的各种人力资源代价总和。

3.6

非人力成本 non-human resource cost

为达成软件测试项目目标所需付出的人力成本之外的其他成本。

3.7

功能点 function point

衡量软件功能规模的一种单位。

3.8

成本度量 cost measurement

在本标准中，特指对软件测试成本的预计值进行估算的过程。

4 软件测试成本构成

4.1 概述

参照标准SJ/T 11463-2013，软件测试成本由直接成本和间接成本构成，如表1所示。

表1 软件测试成本构成

软件测试成本	直接成本	直接人力成本	
		直接非人力成本	工具成本
			环境成本
			其他直接非人力成本
	间接成本	间接人力成本	
		间接非人力成本	

4.2 直接成本

4.2.1 直接人力成本

软件测试直接人力成本包括但不限于测试项目组成员的工资、福利、保险等人力资源费用。软件测试直接人力成本记为 LC 。

软件测试项目组成员包括参与项目测试过程的所有测试或支持人员。根据GB/T 15532-2008的要求，一般情况下，软件测试的人员配备如表2所示。

表2 软件测试人员配备

工作角色	具体职责
测试项目负责人	管理监督测试项目，提供技术指导，获取适当的资源，制定基线，技术协调，负责项目的安全保密和质量管理
测试分析员	确定测试计划、测试内容、测试方法、测试数据生成方法、测试环境（软/硬）、测试工具、评价测试工作的有效性
测试设计员	设计测试用例，确定测试用例的优先级，建立测试环境
测试程序员	编写测试辅助软件
测试员	执行测试、记录测试结果
测试系统管理员	对测试环境和资产进行管理和维护
配置管理员	设置、管理和维护测试配置管理数据库

注1：当开发方实施测试时，配置管理员由软件开发项目的配置管理员承担；当独立的测试组织实施测试时，应配备测试活动的配置管理员。

注2：一人可承担多个角色的工作，一个角色可由多个人承担。

4.2.2 直接非人力成本

4.2.2.1 概述

软件测试直接非人力成本包括但不限于测试环境成本、测试工具成本、其他直接非人力成本等非人力资源费用。软件测试直接非人力成本记为 NL 。

4.2.2.2 测试环境成本

测试环境成本包括但不限于下列成本：

- a) 测试所需要的硬件环境的成本；
- b) 测试所需要的软件环境的成本；
- c) 开发测试所需要的硬件环境的成本；
- d) 开发测试所需要的软件环境的成本。

测试环境成本记为 EC 。

4.2.2.3 测试工具成本

测试工具成本是为测试某个项目而购买或租赁测试工具的费用,以及在测试某个项目的过程中使用已有测试工具的折旧费用和维护费用。测试工具成本记为 IC ,包括自有工具成本和(或)租借工具成本。

自有工具成本记为 OT ,按照具体的测试工期分摊每年的折旧费和维护费,每年的工作日数量按照国家法律法规的规定计算。

租借工具成本为 RT ,按照具体租借费用计算。

4.2.2.4 其他直接非人力成本

直接服务于软件测试项目的其他非人力成本包括差旅费、培训费、业务费等。其他直接非人力成本标记为 OD 。

- a) 差旅费,即为测试某个项目而产生的差旅费用,如交通、住宿、差旅补贴等;
- b) 培训费,即为测试某个项目而安排的特别培训产生的费用;
- c) 业务费,即为完成某个项目测试工作所需辅助活动产生的费用,如招待费、评审费、验收费等;
- d) 采购费,即为测试该项而需特殊采购测试设备之外的专用资产或服务的费用,如专用技术协作费、专利费等;
- e) 其他,即未在以上项目列出但确系为测试某个项目所需花费的费用。

4.3 间接成本

4.3.1 间接人力成本

间接人力成本指服务于测试管理整体需求的非测试项目组人员的人力资源费用分摊,包括但不限于管理人员、组织级质量保证人员、组织级配置管理人员等的工资、福利、保险等的分摊。间接人力成本标记为 IL 。

4.3.2 间接非人力成本

间接非人力成本指不为测试某个特定项目而产生,但服务于整体测试活动的非人力成本分摊。间接非人力成本包括但不限于软件测评资质维护费用的分摊、办公费的分摊。间接非人力成本记为 IN 。

办公费,即为测试某个项目而产生的行政办公费用,如测试场地房租、水电费、物业费、办公用品、通讯、邮寄、印刷、会议及各种办公设备的租赁、维修、折旧分摊等。

注:办公费可以包括以上所述的各项产生的费用,但不限于以上几种,本标准的使用者可以对本标准中列出的各项进行增加、删除和修改。

5 软件测试成本调整因子

5.1 概述

影响软件测试直接人力成本的因素包括但不限于被测软件复杂度、被测软件的完整性级别、测试的质量特性选项、测试的风险度、回归测试、加急测试、现场测试和测评机构资质,不同的因素赋予不同的调整因子。

5.2 被测软件的复杂度

软件的复杂度的度量项包括但不限于算法、数据关系、代码结构、性能内存需求、安全性、保证书、外部接口、兼容性、软件语种等。

被测软件的复杂度可按照以下特性来进行度量：

- a) 算法：极复杂的嵌套和嵌入算法、复杂的微分议程，逻辑决策贯穿始终；
- b) 数据关系：具有主属性及关联关系的多维关系型数据，在多种错误条件下导致的复杂编辑和验证；处理逻辑包括极复杂的数据和数据关系，根据不同的处理逻辑进行选择输入或连续输入；
- c) 代码结构：源代码非常难于阅读和维护；
- d) 性能内存需求：在线、持续可用、严格计时的软件或实时软件，对内存要求很高或有严格要求；
- e) 安全性：包含严格的安全限制，维护关于访问应用程序的人员及其访问的特定数据的设计数据，包括日期、时间和标识；用于涉密机构并被划分为机密级的软件；用于涉密机构的高度敏感的且极端机密的软件；
- f) 保证书：软件可能给用户造成严重损失，软件开发人员将按照合同规定的质量等级受到严厉惩罚；软件运行失败很可能会造成个人或机构的人员伤亡或损失；
- g) 外部接口数量：存在大量外部接口；
- h) 兼容性要求：有特殊的兼容性要求；
- i) 软件语种：非简体中文软件。

软件的复杂度的调整因子为*C*，取值范围如表3所示。

表3 软件复杂性调整因子

复杂性程度	描述	调整因子
A 级	出现上述特性中的四个以上。	1.6~1.8
B 级	出现上述特性中的两个至四个。	1.3~1.5
C 级	出现上述特性中的一个。	1.1~1.2
D 级	没有出现上述任何一个特性。	1.0

5.3 被测软件的完整性级别

软件的完整性级别是系统完整性级别在包含软件部件或者仅仅包含软件部件或仅仅包含软件的子系统上的分别。不同完整性级别软件的安全性要求不同，对软件的测试内容、测试要求和测试所采用方法也有所不同。软件的完整性级别同赋给与软件相关的威胁最高风险类相对应。GB/T 18492-2001 给出了风险级别和系统完整性的映射关系，将软件完整性级别分为A、B、C、D四个等级，如表4所示。

表4 风险等级映射到系统的完整性级别

风险等级	完整性级别
高	A 级
中	B 级
低	C 级
微小	D 级

软件完整性调整因子记为 I ，取值范围如表5所示。

表5 软件完整性调整因子

完整性级别	调整因子
A 级	1.6~1.8
B 级	1.3~1.5
C 级	1.1~1.2
D 级	1.0

5.4 测试质量特性选项

根据GB/T 16260.1—2006的要求，软件质量模型包括功能性（适合性、准确性、互操作性、安全保密性、功能性的依从性）、可靠性（成熟性、容错性、易恢复性、可靠性的依从性）、易用性（易理解性、易学性、易操作性、吸引力、易用性的依从性）、效率（时间特性、资源利用性、效率的依从性）、维护性（易分析性、易改变性、稳定性、易测试性、维护性的依从性）、可移植性（适应性、易安装性、共存性、易替换性、可移植性的依从性）六大质量特性，27个子特性。依据委托方对软件质量特性的测试要求，测试质量特性调整因子记为 Q ，取值范围如表6所示。

表6 测试质量特性调整因子

质量特性	调整因子
功能性（不含安全保密性）	1.0
其他指定质量要求	1.1~1.2
安全保密性	1.3~1.5
效率	1.6~1.8

5.5 测试风险度

软件本身的复杂性以及测试的特点决定了测试活动中风险的大量存在，可能的测试风险由以下部分构成：

- a) 对被测试软件的领域不熟悉；
- b) 被测试软件的开发过程不规范；
- c) 被测试软件出现问题将造成严重后果（例如造成生命、财产损失等）；
- d) 测试时间不足。

测试风险调整因子记为 R ，取值范围如表7所示。

表7 测试风险调整因子

风险程度	描述	调整因子
低	没有出现上述任何一个特性。	1.0
中	出现上述特性 a)、b)、d) 中的一个。	1.1~1.2
高	出现上述特性 c)，或者 a)、b)、d) 中的两个或两个以上。	1.3~1.5

5.6 回归测试

回归测试是指修改了旧代码以后，重新进行测试以确认修改没有引入新的错误或导致其他代码发生错误。回归调整因子记为 I_r ，可取值0.6~0.8。

5.7 加急测试

加急测试可能引发增加测试人员或测试人员加班等现象，从而导致测试成本增加。加急调整因子记为 U ，依据加急程度可取值1.2~3.0。

5.8 现场测试

现场测试由于测试环境离开了测试实验室会导致额外的成本增加。现场测试的调整因子记为 X ，可根据实际情况取值：1.0~1.3。

5.9 评测机构资质

依据评测机构所获得的国家标准、地方标准、行业标准以及机构内外部资质审核等进行调整。评测机构资质调整因子记为 A ，可取值1.0~1.2。

6 软件测试成本度量

6.1 概述

软件测试成本度量分为五个步骤，一是进行软件规模度量，二是根据软件规模估算软件测试的工作量，三是根据工作量估算软件测试工期，四是根据工期估算软件测试各部分成本，五是根据估算的各部分成本确定软件测试的成本。软件测试成本估算的基本流程，如图1所示。

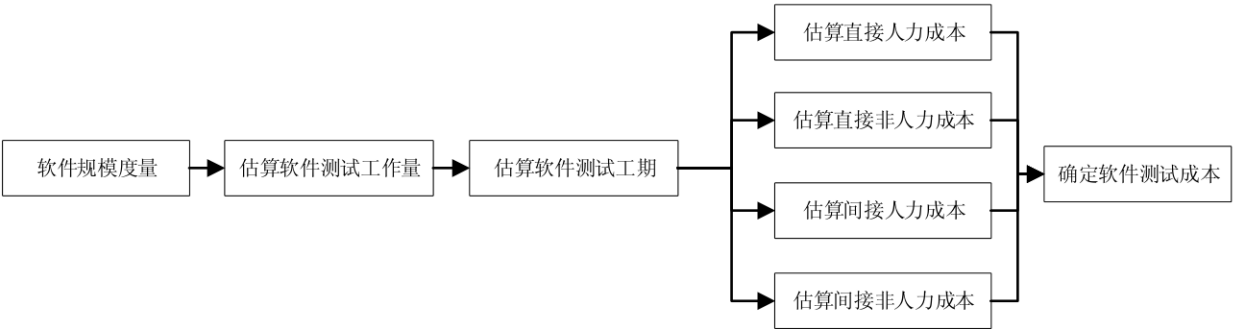


图1 软件测试成本估算基本流程

6.2 软件测试成本估算基本流程软件规模度量

在软件规模度量前，应根据测试计划或测试合同中规定的测试范围确定系统边界，然后再根据已确定的系统边界和需求描述、项目特点和度量需求进行软件功能规模度量。在软件规模度量时，应考虑可能的测试需求变更程度，并对规模度量结果进行适当调整。

软件规模度量的方法可选用国际标准化组织ISO/IEC已发布的以下五种功能规模度量标准中的一种，即：

- a) ISO/IEC 19761 (COSMIC-FFP 方法)；
- b) ISO/IEC 20926 (IFPUG 方法)；
- c) ISO/IEC 20968 (Mk II 方法)；
- d) ISO/IEC 24570 (NESMA 方法)；
- e) ISO/IEC 29881 (FiSMA 方法)。

注：根据相关国际标准中的方法适用范围声明，COSMIC方法适用于商业应用软件和实时系统；IFPUG方法适用于所

有类型软件的功能规模度量；Mk II 方法适用于逻辑事务能被确定的任何软件类型；NESMA方法与IFPUG方法非常类似，但对功能点计数进行了分级，以便在估算的不同时期选择不同精度的方法进行估算；FiSMA方法适用于所有类型软件的功能规模度量。

6.3 估算工作量

6.3.1 概述

工作量估算分为估算准备、估算与调整两步。

6.3.2 估算准备

估算准备主要是分析测试的风险、确定影响测试的主要因素。在进行工作量估算前，应：

- a) 对测试项目风险进行充分分析。风险分析时应考虑技术、管理、资源、商业多方面因素。例如：需求变更、外部协作、时间或成本约束、人力资源、系统架构、用户接口、外购或复用、采用新技术等。
- b) 根据经验或相关性分析结果，确定影响工作量的主要属性。
 - 1) 软件复杂度；
 - 2) 应用领域；
 - 3) 质量要求，如可靠性、可使用性、效率、可维护性、可移植性等；
 - 4) 采用技术，如开发平台、编程语言、系统架构、操作系统、测试工具软件等；
 - 5) 测试团队，如测试方组织类型、团队规模、人员能力等；
 - 6) 过程能力，如测试方过程成熟度水平、管理要求等。
- c) 选择合适的工作量估算方法。

6.3.3 估算与调整

估算与调整主要是根据风险分析结果，对估算方法或模型合理调整，根据软件测试复用情况的分析调整对工作量的估算。

在进行工作量估算时，应根据本标准第5章所述的调整因子进行工作量调整。

宜采用不同的方法分别估算工作量并进行交叉验证。如果不同方法的估算结果产生较大差异，可采用专家评审方法确定估算结果，也可使用较简单的加权平均方法。

在估算工作量时，宜给出估算结果的范围而不是单一的值。

根据GB/T 25000.51的要求，软件测试包括产品说明评审、用户文档集评审、软件质量测试、测试文档集编制和符合性评价五部分。

软件测试工作量可按照以下两种方式进行度量：

- a) 依据软件测试生存周期，按照各个阶段计算测试（含评审）工作量。软件测试生存周期包含以下阶段：
 - 1) 测试需求：根据软件需求规格说明确定测试需求；
 - 2) 测试策划：确定测试的内容或质量特性，确定测试的充分性要求；
 - 3) 测试设计：根据测试要求、被测软件文档和测试规范，确定测试的方法；
 - 4) 测试环境准备：准备测试需求的各种环境，测试代码开发，包括设计驱动模块和桩模块；
 - 5) 测试数据准备：测试执行前，为测试能够进行准备一组可以验证的数据；
 - 6) 测试用例开发：测试用例设计，包括自动化测试时录制和编辑测试脚本；
 - 7) 测试执行：依据测试用例执行软件测试，并记录测试结果，包括手工测试和/或自动化测试；

- 8) 测试结果分析：对测试结果进行分析；
 - 9) 测试报告编制；
 - 10) 测试评估：对测试进行分析及评审，评价软件符合性，包括测试的收益、软件将来可能存在的风险。
- b) 依据功能点法进行软件规模度量。

针对最终计算出的功能点数目，按照“K人日/功能点”来计算测试工作量。

注：软件规模度量过程中所产生的费用应由委托方承担。

6.4 估算工期

对软件测试工期的估算主要有以下几种方式：

- a) 根据工作量估算结果和资源情况，对软件测试任务进行分解并制订时间表；
- b) 利用基准数据建立“工作量-工期”模型，估算合理的工期范围；
- c) 将委托方的期望工期或初步制订的工作时间表中的工期与工期估算结果进行比较；如果委托方期望工期或工作时间表中的工期短于估算出的工期下限或超出估算出的工期上限时，应分析原因，必要时需对人力资源安排或项目范围进行调整，再重新估算工作量、工期，并制订新的工作时间表。

6.5 估算直接人力成本

根据工作量估算结果和测试人员直接人力成本费率估算直接人力成本。应优先使用本组织的直接人力成本费率数据。直接人力成本的计算宜采用以下两种方式：一是根据不同类别人员的直接人力成本费率和估算工作量分别计算每类人员的直接人力成本，将各类人员的直接人力成本相加得到某个项目的直接人力成本；二是根据项目平均直接人力成本费率和估算的总工作量直接计算某个项目的直接人力成本。

对于非全职投入软件测试工作的人员，按照项目工作量占其总工作量比例折算其人力资源费用。

6.6 估算直接非人力成本

根据项目情况，按照第4章中的要求分项估算直接非人力成本。

6.7 估算间接人力成本

估算人员根据项目情况，按照第4章中的要求分项估算间接人力成本。间接人力成本应按照工作量比例进行分摊。

6.8 估算间接非人力成本

估算人员根据项目情况，按照第4章中的要求分项估算间接非人力成本。间接非人力成本应按照工作量比例进行分摊。

6.9 度量公式

软件测试成本度量可按如下公式计算：

- a) 软件测试工作量调整：

$$LW = UW \times (C + I + Q + R + U + X + A) \times (1 + n \times T_r) \dots\dots\dots (1)$$

式中：

LW——调整后的软件测试工作量，单位为人日（pd）；

LW——未调整的软件测试工作量，单位为人日（pd）；
C——复杂性调整因子，取值范围1.0~1.8；
I——完整性调整因子，取值范围1.0~1.8；
Q——测试质量特性调整因子，取值范围1.0~1.8；
R——风险调整因子，取值范围1.0~1.5；
U——加急调整因子，取值范围1.2~3.0；
X——现场测试，取值范围1.0~1.3；
A——评测机构资质调整因子，取值范围1.0~1.2；
T_r——回归调整因子，取值范围0.6~0.8；
n——回归测试次数。

注：因项目变化导致需要重新进行工作量估算时，应根据该变化的影响范围对工作量估算方法及估算结果进行合理调整。

b) 软件测试直接人力成本计算：

$$LC = LW \times S \dots\dots\dots (2)$$

式中：

LC——软件测试人力成本，单位为元；
LW——软件测试人力成本工作量，单位为人日（pd）；
S——工作量单价，依据当地平均收入水平调整，单位为“元/人日”。

c) 测试工具成本计算：

$$IC = OT + RT \dots\dots\dots (3)$$

式中：

IC——软件测试工具成本，单位为元；
OT——自有工具成本，参照本标准4.2.2.3节计算，单位为元；
RT——租借工具成本，依据租借费用估计，单位为元。

d) 软件测试直接非人力成本计算：

$$NL = EC + IC + OD \dots\dots\dots (4)$$

式中：

EC——软件测试环境成本，单位为元；
IC——软件测试工具成本，单位为元；
OD——软件测试其他直接非人力成本，单位为元。

e) 计算软件测试成本：

$$STC = LC + NL + IL + IN \dots\dots\dots (5)$$

式中：

STC——软件测试成本，单位为元；
LC——软件测试直接人力成本，单位为元；
NL——软件测试直接非人力成本，单位为元；
IL——软件测试间接人力成本，单位为元；
IN——软件测试间接非人力成本，单位为元。