密级：开源

学习曲线组件设计

OFBiz-LearningCurve 组件

北京朗华世纪科技发展有限公司
WWW.LANGHUA.CN
2008 年 11 月

目的和需求

在公司自用的项目管理系统中，需要对与软件开发中的生产活动进行预期，例如，开发一个全文检索模块，张三做会用多少时间？李四做会用多少时间？

学习曲线是做这种预期的方法之一。本组件即是计算期望值的算法组件，应实现以下需求：

1. 能够根据输入的用户信息和功能点列表，计算出活动预计所需时长。
2. 能够根据预计所需时长以及报警线百分比，计算出提前完成报警时间和延迟完成/未完成报警时间。

实现原理

“学习曲线”概念请参考 http://en.wikipedia.org/wiki/Experience_curve_effects。

根据学习曲线理论，可以形成代价（费用、成本、时间等）与做事数量的计算公式。美国 NASA 和美国联邦航空管理局有这方面的算法说明：

- http://cost.jsc.nasa.gov/learn.html
- http://fast.faa.gov/pricing/98-30c18.htm

下面对这两个网页的翻译，以期开发人员对这个原理能够有较为深入的理解。

另外，从历史数据得到学习曲线公式，在数学上使用的是一元线性回归方法，对这个方法的说明见：

- http://en.wikipedia.org/wiki/Linear_regression
- http://en.wikipedia.org/wiki/Least_squares

NASA 的学习曲线计算器

在 1936 年 2 月的《航空科学》杂志发表的一篇文章中，T.P. 怀特为航空工业引入了学习曲线概念。怀特描述了获得基于飞机组装这种重复生产的成本估算的基本理论。从此，学习曲线在各类工作中得到应用，从简单任务到复杂的航天飞机制造。

学习曲线理论很简单。它认为重复同一个操作会导致使用更少的时间或更小的努力。在怀特的学习曲线中，隐含的假设是完成一个单元的生产所需的直接劳动工时数，会在每次生产数量加倍时，按一个固定的百分比下降。如果在数量加倍时改进的比率是 20%，那么“学习百分比”会是 80%(100-20=80)。虽然学习曲线强调时间，它也能很容易地扩展为成本。

学习百分比通常由对类似产品的实际成本数据统计分析确定。如果缺少这类数据，你可能要使用 Rodney Stewart 编写的《Cost Estimator's Reference Manual- 2nd Ed.》中的指导原则：

- 75%的手工组装/25%的机加工 = 80%
- 50%的手工组装/50%的机加工 = 85%
- 25%的手工组装/75%的机加工 = 90%

或者

1. 航空 85%
2. 造船 80-85%
3. 新型复杂机械工具 75-85%
4. 重复的电器制造 90-95%

5. 重复的机加工或冲压操作　90-95%
6. 重复的电器操作　75-85%
7. 重复的焊接操作　90%
8. 原料　93-96%
9. 采购零部件　85-88%

这里的计算器使用学习曲线来预计生产单元达到指定数量的单个、平均和全部投入。投入可以用成本、工时或任何其它度量单位来表示。这里的计算器能够计算怀特学习曲线或 Crawford 学习曲线。用户输入生产第一个单元的投入（成本、工时等）、总数、学习百分比。

对于这些计算学习曲线数值的方法的详细解释，包含在了 Phillip F. Ostwald 所写的《Engineering Cost Estimating》。你还能在 Matthew S. Goldberg 和 Anduin E. Touw 所著的《Statistical Methods for Learning Curves and Cost Analysis》找到这些方法。

美国联邦航空管理局对学习曲线的说明

第十八章　学习曲线

1. 简介

在第十五章数量分析技术中介绍了学习曲线的概念。（成本改进法也引用了学习曲线理论。）学习曲线用在了成本估算和分析中。估算侧重在得到学习曲线，分析则强调对分包商提供的学习曲线进行检查。基本概念是相同的，只是应用上有差别。

本章会较为详细地浏览学习曲线理论，并聚焦在它的应用上。第二节会解释学习曲线的基本概念。第三、四节描述了通用和专用的学习曲线应用。第五节讨论评价学习曲线时使用的标准。

2. 基本概念

学习曲线来自从历史角度观察个人从事重复任务时所表现出的任务重复次数与能力进步之间的关系。对这个现象的经验研究为现在的理论和实践产生了三个结论：

- 执行一个任务所需要的时间随着任务的不断重复而减少
- 做得次数越多，继续改进的量越小
- 改进的比率在足够大的范围内是一个常量，可以使用它来作为预测工具

研究发现，改进常量是生产数量加倍时所需时间按一个常量百分比减少。这个随数量加倍而减少的常量百分比称为“学习比率”。学习曲线的斜率是 100 减去学习比率。例如，如果数量加倍时工时减少 20%（学习比率），那么学习曲线的斜率是 80%。

当在坐标纸上绘图时，这条线会变成一条双曲线。这是因为成本降低的数量不是一个常量。而且，随着数量加倍，减少的数量在逐渐减少。当数量很大时，减少量会变得可以忽略不计了。这发生在工作人员接近一个工艺的“标准成本”时。在工艺或产品持续变化时，这个点会延迟。在钢铁生产、汽车组装、服装和乐器行业的多个研究把这个现象描述为“停滞”。

尽管随着时间成本实际下降的量也在下降，但是成本在加倍期间改变的量在实验研究中是一个常量百分比（T.P. 怀特，波士顿咨询）。咨询顾问们简化了这个估算。这个情况可以表示为一个指数等式。指数等式可以用对数来求解。在以成本估算交易时，常常会提到在对数纸上画学习曲线。今天，计算机取代了对数纸。

3. 单元和累计平均理论

学习曲线明确的类型（如数学模型）会经常以提议人或率先使用的公司来命名，比如怀特、克劳福德、波音、诺斯罗普。所有这些名字都是指两个公认的、最佳描述了随生产数量增加而成本或工时减少的数学模型之一。这两个模型是指单元曲线或累计平均曲线。下面的模型公式看起来一样。然而，由于相关的变量定义不同，即使使用相同的第一个单元（又称为理论单元 1，简称 T1）和斜率值，它们预测的结果也会不同。因此，在评估用来预测成本的斜率之前，分析人员必须注意斜率是用于单元法的还是累计平均法的。

下面两个等式仅在 Y 的定义上有区别，但是这个区别会对生成的估算产生显著差别。

等式 18-1 描述了单元曲线，Y 表示在这一轮生产中指定单元的成本。如果一轮生产要制造 200 个单元，那么全部成本能够通过运行这个等式 200 次，分别算出单元 1 到 200，然后把这 200 个数值累加起来。

Equation 18-1. Unit Theory

$$Y_X = T1 \bullet X^b$$

Where:
Y_X = The cost required to produce the X^{th} unit
T_1 = The theoretical cost of the first production unit
X = The sequential number of the unit for which the cost is to be computed
b = A constant reflecting the rate costs decrease from unit to unit

等式 18-2 描述了累计平均曲线。在这个等式中，Y 代表不同数量（X）的单元的平均成本。在累计平均中累计的含义是计算 X 个累计单元的平均成本。因此，X 个单元的全部成本是 X 乘以累计平均成本。例如，要计算单元 1 到 200 的全部成本，分析人员可以计算单元 200 的累计平均成本，然后用这个值乘以 200。这个方法比单元曲线法容易很多。

Equation 18-2. Cum Average Theory

$$Y_X = T1 \bullet X^b$$

Where:
Y_X = The average cost of the first X units
T_1 = The theoretical cost of the first production unit
X = The sequential number of the last unit in the quantity for which the average cost is to be computed
b = A constant reflecting the rate costs decrease from unit to unit

在这两个学习曲线之间还有很重要的相似之处。两者都是实验功能。这是个重要信息，因为解决实验等式需要使用对数。根据 Donald Stancl 在《管理、生活和社会科学》中的论述，“解决方法是对等式两边取对数，然后使用 $\ln X^b = b*\ln X$ 把 b 从指数中取出来。”方便起见，通常使用自然对数（在数学书和电子表格中常用 ln 表示）。了解使用对数解学习曲线等式，会帮助读者有效使用电子表格工具。

对等式两边取自然对数，等式变成数学上的直线方程：

lnY=T1+b*lnX

或

Y=aX+b

直线对分析人员很有用，因为直线容易扩展到数据范围之外——如果用渐近线来寻找一组数据的最小平方，扩展直线和从直线来估计要容易一些。相反，对于曲线，一组数据之外的精确角度较难推导出来。总之，累计平均学习曲线可以写成下面的等式：

Equation 18-3 Logarithmic Transformation of the Learning Curve Equation

$$\ln(Y) = T1 + b \bullet \ln(X)$$

Where:

Y_x = *The average cost of the first X units*

T_1 = *The theoretical cost of the first production unit*

X = *The sequential number of the last unit in the quantity for which the average cost is to be computed*

b = *A constant reflecting the rate costs decrease from unit to unit.*

$\ln$ = *The natural logarithm*

单元和累计平均学习曲线等式都对数量加倍成本按一个常量百分比下降进行了描述和建模。这个常量百分比是通过 b 值反应出的。两个等式的 b 值可以通过下面的等式计算：

Equation 18-4 Value of b

$$b = \ln S / \ln 2$$

Where:

S = *The cost/quantity slope expressed as a decimal value.*

举个例子，如果第一个单元成本是 100，第二个单元成本是 90，或第一个单元的 90%，那么单元曲线会是 90%的斜率，S 值会是 0.9。b 值会是(log 0.9)/(log 2)或-0.045758/0.30103 或-0.15200。当然，这个数据集合很小，读者应该知道，两个数据点不会产生一个可信的预估值。这里只是举个例子来说明 b 是如何与学习曲线的斜率相关的。在谈及学习曲线时，更常会讨论的是百分比的斜率，如 90%。当从标准的电子表格学习曲线等式分析工具中得出学习曲线等式时，通常这类工具只会给出 b 值。因此，分析人员必须理解如果把 b 值转换为斜率。当然，这可以通过下面的公式得出：

Equation 18-5 Equation to Determine Slope from b

$$\ln \text{slope} = b * \ln 2 \text{ or}$$

$$e^{b*\ln 2} = \text{slope}$$

使用 T1 值(100)和从上面的单元曲线样例得到的斜率(0.9)，累计平均曲线总是会比单元曲线得到一个更低的成本合计，因为差异出在了 Yx 是如何定义的。在上面的例子中，第一个单元成本 100，第二个 90，按单元曲线法，两个合计是 190。如果使用累计平均曲线，同样的 T1 值(100)，斜率是 0.9，两个单位的全部花费是 2 倍的 90 或 180。

3.1 学习曲线计算中的典型问题

本节描述了成本估算所需的学习曲线的一些典型情况或问题。这些情况在下面使用对数比例画成本－数量时阐述。现在讨论转向了如何进行计算以及使用什么工具来进行估算。每个问题会作为一个案例来讨论。

第一个问题是从一组原始数据计算实际学习曲线参数（T1 和斜率），发展出一个学习曲线模型。这个问题是案例一。本节中的其它问题则是对一个已知的学习曲线（由斜率和 T1 定义）的应用。

案例一：第一部分——介绍：从数据得到 T1 和斜率数值

背景：

假定使用下列实际数据作为预测的基础：

单元数据

单元序号	单元成本
1	100
2	96
3	88
4	83
5	78

- 两个或更多单位成本值，或
- 两份或更多份成本

简单数据绘图

目标：

计算 T1 和斜率 b 数值。

方法：

首先，分析人员必须确信收集到的只是重复发生的数据——非重复发生的数据不应该包括在数据集合中，因为它不会与学习相关。必须完成了其它数据标准化，比如剔除通胀因素。而且，分析人员必须决定这些数据是否要转换为对数。如果使用学习曲线软件，这个转换通常是由软件完成的，对用户透明。如果用户倾向使用标准电子表格渐进分析，这些数据会被转换成对数。本节避免讨论某些学习曲线软件，而是聚焦在使用标准的电子表格来计算学习曲线。大量的软件都内置了学习曲线工具。某些软件包只是做学习曲线分析，而其它内置的学习曲线工具具有估算能力。在互联网上搜索能找到一个早期的（CURV1 Learning Curve Analysis for Windows）和后期的（ACEIT）例子。

分析人员应该知道这些程序是如何工作的，以便能在电子表格中使用标准的渐进分析工具，或者手工计算。手工计算也是对计算机生成结果的一个好的校对检查。请记住，错误的输入会产生错误的结果。

所有学习曲线程序会通过渐进分析法从数据计算 T1 和斜率，形成符合数据集合的一条直线。这些程序常用的步骤会以商业电子表格程序中使用的渐进分析工具为例来进行说明。大多数电子表格都把渐进分析工具作为他们程序的一个标准组成部分。这些工具会要求分析人员在电子表格中输入数据，然后告诉工具数据在哪里。今天，这些工具对用户非常友好，会在处理过程中提示用户。然后电子表格渐进分析工具会根据这些数据计算斜率和 T1。大

多数应用还提供基于计算出来的 T1 和斜率得到 X 个数值的选项。

每个软件都会有所不同，因此最好是阅读用户手册或帮助。如果分析人员不能让程序运行起来，可能会是输入错误造成的。软件厂商应该会在这方面提供帮助。

案例一：第二部分——从单元数据得到斜率和 T1

背景：

如果有历史单元数据，这些数据可以直接使用。分析人员必须确信这些数据是重复发生的。非重复发生的数据没有学习效果。

因为这个例子使用了标准的电子表格渐进分析工具，必须把数据转换为对数。对数转换和渐进分析工具的输出如下图所示。典型的对数转换可以通过在电子表格的单元格中输入 ln(变量)来实现，变量表示了数据中的每个值。然后那些单元格成为渐进分析工具的输入。这个工具可以计算一些标准的渐进统计，帮助确定学习曲线的使用情况以及学习曲线等式的系数。本例中电子表格计算的节点是自然对数表，因为 X 和 Y 数据输入到对数表中了。因此，节点值必须进行转换以便表现为标准的学习曲线等式 $Y=T1*X^b$ 格式。这是由电子表格来完成的。b 值不需要做转换。因此，学习曲线等式会是 $Y=102.73*X^{-0.15}$。直线的斜率是 $e^{b*\ln(2)}$或 0.90。这个数据集合，按照最小方差，确定是 90%的学习曲线。

解题：

单元数据

单元序号	单元成本
1	100
2	96
3	88
4	83
5	78

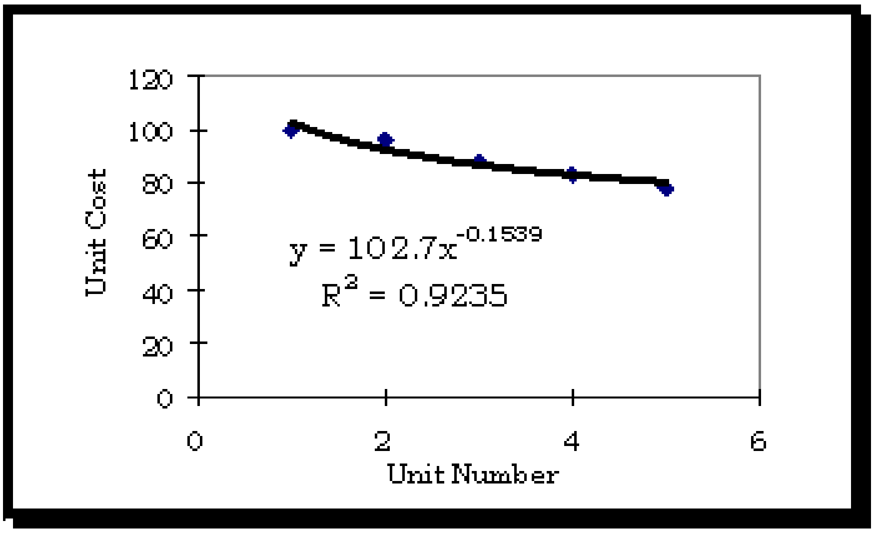
输出：

T1: 102.7

b: -0.1539

斜率: 0.90

R^2: 0.9235



单元学习曲线

对数转化后的单元数据

ln(X)	ln(Y)
0	4.605
0.693	4.564
1.099	4.477
1.386	4.419
1.609	4.357

对数线性学习曲线等式：ln(Y)=T1+b*ln(X)

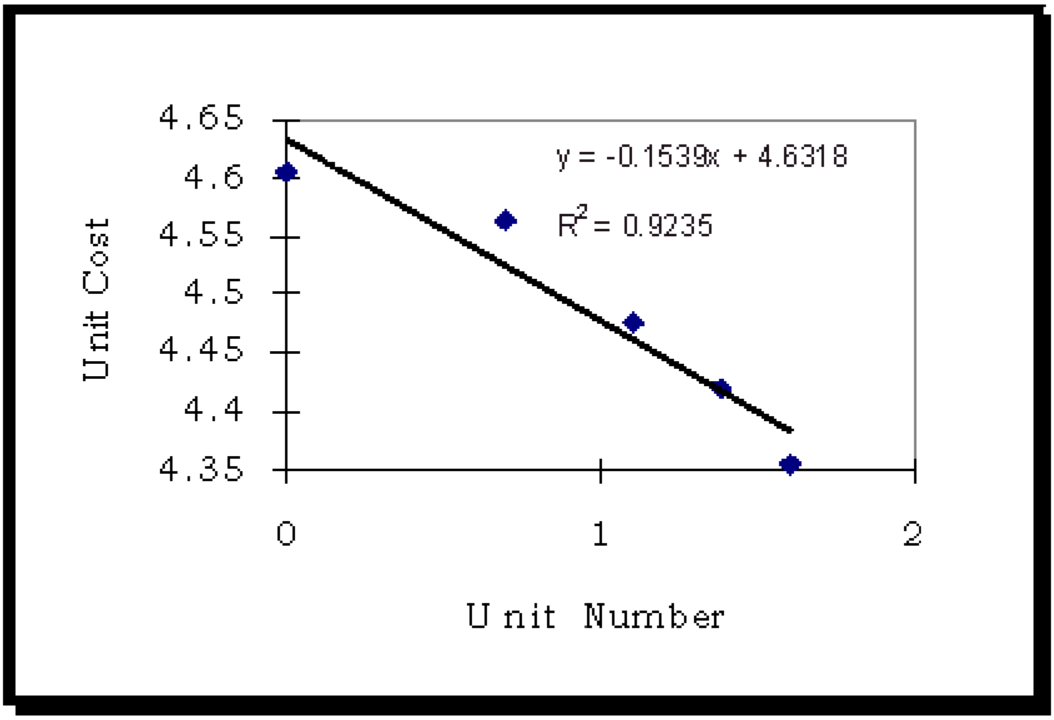
输出：

T1=exp(4.6318):102.7

b: -0.1539

斜率: 0.90

R^2: 0.9235



对数单元学习曲线

案例一：第三部分——从份数数据得到斜率和 T1
背景：

使用份数数据时，分析人员需要在把数据输入到软件工具之前做一些数据转换。有些学习曲线程序会自动做转换，计算"真正份数绘图点"。分析人员应该查阅软件的用户手册以保证正确使用工具。下面的讨论会让读者理解对这些份数数据需要进行转换。

使用份数数据时，分析人员通常只知道一份中单元的数量以及一份的成本合计。因为学习曲线理论假设在生产一定数量的单元（而不是份数）时，成本会按一个常量百分比变化，所以必须把一份的成本转换为一个单元的成本。计算机程序已经开发出来了为学习曲线计算"真正份数绘图点"。如果必须手工调整中点，下面的指南可以用来估计一个"真正份数绘图"的点。

一份数据绘图点指南

累计平均理论

对于累计平均曲线理论，来自连续份数的数据会被添加到先前的成本和数量中，以便得到新的份数相关的那些绘图点。

份数 Y 轴绘图点：Y 轴绘图点是全部成本除以单元数。
份数 X 轴绘图点：X 轴绘图点是份数中的最后一个单元。

单元曲线理论

份数 Y 轴绘图点：Y 轴绘图点是全部成本除以单元数。
份数 X 轴绘图点：计算 X 轴绘图点时使用了下列约定：

- 如果第一份小于 10 个单元，那么绘图点是中点（单元数除以 2）。
- 如果第一份大于等于 10 个单元，那么绘图点是单元数除以 3。
- 以后各份的绘图点是每份单元数除以 2。

解题：

使用上述指南，累计平均和单元理论的数据集合如下表所示：

份数序号	单元数量	份成本	单元曲线 X 绘图点	单元曲线 X 绘图点	累计平均 X 绘图点	累计平均 Y 绘图点
1	2	200	1	100	2	100
2	3	285	3.5	95	5	97
3	5	450	7.5	90	10	93

在这个样例中，累计平均曲线公式将用来获得 T1 和斜率。因此，我们会使用上表中的最后两列作为 X 和 Y 的数据，输入到电子表格的渐进分析中。把这个插入一个标准的电子表格渐进分析程序，你将用到最小方差工具。首先，你必须把数据取对数；然后输出也必须取对数。

由渐进分析法得出的学习曲线等式是 $Y=103.45*X^{-0.04}$。直线的斜率是 $e^{b*\ln(2)}$或 0.97。这个数据集合，按照最小方差法，是 97%的累计平均学习曲线。

案例二：从单份线性曲线估算

背景：

T1、斜率、曲线类型（单元或累计平均）和份数中的单元数量是估算所必须的。

使用案例一中获得的学习曲线参数。

单元曲线： $\bar{Y}=102.73\cdot X^{-0.15}$

累计平均曲线： $\bar{Y}=103.45\cdot X^{-.04}$

目标：

估算有 5 个单元的份数 4 的成本合计。

使用累计平均理论解题：

前面介绍了累计平均曲线可以由下面的数学等式描述：

Equation 18-6 Mathematical representation of the cum average curve theory.

$$\bar{Y} = T1 \bullet X^{b}$$

Where:

b = log slope /log 2

$\bar{Y}$ = *The average cost of the first X units*

T_1 = *Cost of first unit*

X = *Number of units.*

在案例一中，适用于累计平均曲线的数据可以描述为：

$$\bar{Y}=103.45\cdot X^{-.04}$$

同时，前 X 个单元的累计成本合计由下面的等式给出：

Equation 18-7. Cumulative Total equation.

$$Y_T = T1 \bullet X^{b+1}$$

Where:

Y_T = *the cumulative total lot cost.*

要确定某一份或多份成本，使用上式来计算累计成本。这会产生全部单元的成本合计。第四份有 5 个单元。1-3 份总共有 10 个单元。因此，第四份的最后一个单元是单元 15。

$Y_{15} = 103.45 \times 15^{.96} = 1392$

这 1392 个小时代表了 1-4 份的成本。要获得第四份的成本，需要减去前三份的成本。前三份的单元数是 10，因此：

$Y_{10} = 103.45 \times 10^{.96} = 943$

第四份的成本是

$Y_{11,15} = 1392 - 943 = 449$

使用单元理论解题：

单元理论的等式如下：

Equation 18-8

$$Y_x = T1 \bullet X^b$$

Where:

$b = \dfrac{\log slope}{\log 2}$

Y_x = *The cost of the* x^{th} *unit*

T_1 = *Cost of first unit*

X = *Unit number*

不幸的是，分析人员使用单元曲线理论的话，没有等式可以快速计算全部份数的成本合计。当然，单元成本可以使用上面这个等式来计算，不过，使用计算器计算每个单元的成本然后加和是很困难的。当要生产的份中包含了第一个单元时，下面等式是个很好的估算方法（N 是份中的单元数）：

Equation 18-9 Approximation of Total Lot Cost

$$Y_T = \frac{T_1}{[b + 1]} * \{ [N + 0.5]^{b+1} + [b + 1] - 1.5^{b+1} \}$$

使用份数 1 来计算它的成本：

$Y_{1,2}$ = (102.73/(-.15 + 1)) · $((2 + 0.50)^{-.15+1} + (-.15 + 1) -1.5^{-.15+1})$

$Y_{1,2}$ = 196

回到案例一的原始数据，我们注意到实际成本是 200 小时。估算值(196)有些错误，但是对于快速估算或者复查电子表格/学习曲线软件的计算结果，已经足够了。

当要计算的那一份没有包括第一个单元是，下面的等式是一个很好的估算方法：

Equation 18-10 Approximation of Total Lot Cost

$$Y_{lot} = T_1 \bullet ((X_L + 0.5)^{b+1} - (X_1 - 0.5)^{b+1})/b+1$$

Where:

Y_{lot} = The total lot cost for items starting with L (unit X_1) and ending with the last unit X_L

第四份有 5 个单元。在第四个之前，总共生产了 10 个单元。因此，第四份里的最后一个单元是单元 15、第一个单元是 11。

Y_{lot} = 102.73 $(((15 + 0.5)^{-.15+1} - (11 - 0.5)^{-.15+1})/-.15 + 1)$

= 350

到这里可以结束学习曲线的基础知识的讨论。在进行一个较为复杂的学习曲线应用之前，将会解决使用哪种曲线。下一节会提出一些建议。

3.2 选择累计平均还是单元曲线?

单元学习曲线技术是政府和合同评估人广泛使用的方法论。以下问题限制了累计平均学习曲线理论的使用：

- 如果是一个特定的公司来预计工作，而工厂有自己的方法，那么应该使用承担了最多工作的那个工厂所用的理论。
- 如果没有选定分包商，那么应该使用可能获得合同的大多数公司使用的理论。
- 如果要使用从历史记录得到的斜率数据，那么应该使用对历史记录的斜率数据最佳的理论。
- 如果有使用过去数据的特定程序能支持对进一步生产进行预测，那么应该使用能对过去数据提供最佳对数线性分析的理论。
- 如果涉及了专用的曲线情形（如阶梯函数）的分析，那么单元法可能比累计平均理论更有用。选择单元理论对阶梯函数要合理些，因为这类中间程序的变化更经常地定义为单元成本变化。
- 当完全无法选择时，使用两个理论都估算以下，看看不确定方面包括不确定的斜率等的差异有多大。

本节提供了成本改进现象和两个最常用的对数线性数学模型的深入理解。然而，学习曲线现象在做预估时有时候是不够的。这是因为每个新系统都是由不同的条件和独得性组合而成的，这些可能不是对历史数据的重复。因此，分析人员面临挑战：理解要制造的新系统的环境、预期环境中会发生的成本改进、能表达这个改进的学习曲线的形状和斜率。下一节处理来自对数线性学习曲线样例的一些“异常”或偏差。

4．特殊学习曲线应用

到这里为止，学习曲线都是在连续生产和稳定的产品设计的假设下。有时，分析人员会发现政府有时购买的成品有一个微小的设计变化，这涉及了需要学习。政府也购买新设计和需要开发的物品。在这两种情况下，对数线性学习曲线似乎是一个恰当的预估工具。在这类场景中，分析人员只需要决定一个恰当的斜率和T1。此后的计算会非常简单明了。

然而，经常在现实生活的生产过程中发生大量争执。过程改变可能导致一个物品会比先前生产的经历一个更陡峭的斜率。可能是生产上在物品上做了添加或去除，或者政府在生产过程中引起了一个暂停，或者因为没有足够资金而导致产出率变化了。发生这些事情会导致单元成本不按学习曲线变化。因此，在预估时必须考虑这类变化的破坏。

4.1在处理对数线性学习曲线的中断时的常规思考

通常，把对数线性学习曲线的中断处理为把情况分为一段一段的片段，然后对每个片段分别进行估算。要对预估的情况概念化，分析人员能够创建实际数据的离散图形。当预估涉及到一个新的同样或类似物品组成的份时，会出现这个情况。类似的，分析人员能分析一个要预估的程序，然后找到要使用的、好的相似程序。在任何条件下，分析人员都必须理解造成数据点发散的因素，并从历史记录元素中选择与当前系统最相似的来进行预估。经常出现不存在相似的情况。这时，分析人员必须利用个人知识和经验，同时从他人那里获取经验，来定义最适合的曲线。

无论使用哪种学习曲线进行预测，分析人员都必须准备好对所选曲线、它的形状、斜率以及预期行为进行辩护和证明。下面各节试图提供在这类挑战下的深入思考。这会占去两个问题，每个都会影响学习曲线的形状，通常会让它的形状偏离纯粹的对数线性形状。问题3（在案例三中讨论）解决曲线斜率改变经常会导致分裂的学习曲线的问题。问题4 （案例四）解决从学习曲线路径“阶梯向下”或“阶梯向上”。

案例三：分裂的学习曲线

背景：

给出了T1、原始斜率、新的斜率、代表曲线上单元数的点以及要估算的份中的单元数。假设斜率因为在第四份中添加了一个新的元件而变化了，这个元件要求不同的、更复杂的装配。一个工程预期说新的斜率会是85%。

数字：

使用和案例一的第三部分同样的数据集合，产生的学习曲线等式是$Y= 103.45 * X^{-.04}$。

份序号	单元数	份成本
1	2	200
2	3	285
3	5	450
4	5	

目标：

计算第四份中生产全部单元的成本，使用累计平均理论。

解题：

- 第一步：为曲线的第二部分计算假设的第一个单元成本：

$$T1' = (T1 \cdot n^{b})/n^{b'}$$

这里，

T1' = 用于新的曲线斜率 S'的假设的第一个单元成本

T1 = 原来的第一个单元成本

n = 发生斜率变化的那个单元数

b = LogS/Log2

b' = LogS'/Log2（新斜率的 b）

S = 原来的斜率

S' = 新的斜率

T1' = (103.45*11-0.04)/11-0.23447 = 165

- 第二步：为第四份（在生产次序上是单元 11-15）计算份成本，使用 T1=164、斜率=85%。累计平均份成本为：

$$Y_T = T1 \cdot X^{b+1}$$

使用这个公式获得四份产生的累计成本合计：

$$Y_{1,15} = 165 \cdot 15^{.765535}$$

$$= 1312$$

使用同样的公式得到前三份的累计全部成本：

$$Y_{1,10} = 165 \cdot 10.765535$$

$$= 962$$

上面两个数字的差是第四份的成本：

$$Y_{11,15} = 1312 - 962$$

$$= 360$$

斜率会发生变化的情况：

很多因素会导致斜率变化。在决定使用那个斜率时应考虑下列因素：

- 新的物品和以前生产的物品的相似性；
- 工艺和元件的增加和减少；
- 材料的差异；
- 以前生产的物品的工程变化的影响；
- 生产一个类似物品的用时；
- 工具和设备的状况；
- 人员利用率；
- 工作条件或士气变化；
- 类似物品之间的其它可比因素；
- 生产效率；
- 原料和元件充足；
- 比较实际生产数据与之前推测或理论曲线的偏差情况。

上述因素会导致斜率变化，也能导致理论的第一个单元成本变化。换句话说，把单

元成本向上或向下调整会更合适，例如去掉或增加一个元件的情况。如果工艺没有很大变化，而去掉了一个元件，可能不会造成学习曲线斜率变化，但是应该调整 T1。技术专家（行业工程师或其他工程师更容易理解预估涉及的工艺）应该帮助决定是否调整斜率或 T1。分析人员负责提出正确的问题，比如上面描述的那些，然后会发现技术专家很乐于对学习曲线的适当调整提供有用的输入。

变更（元件的增加、去除、替换）可能会表现为原来工作量的百分比，作为工作量百分比，或工时数，或钱数。如何应对变化的这个方法说明变化是分散的（如增加或删除一个元件如雷达）。然而，全部变更不都是这么简单明了。相反，变更可能是：元件、支撑结构、外围设备从一个地方移动到了另一个地方；重新布线；因设计变更而导致额外的钻孔或加工等等。因此，以什么形式在学习曲线里表现变化需要根据变化本身的特性来决定。对于更为复杂的变更，可能是斜率变化，因为通常工艺越复杂，斜率越陡。然而，下一个问题中的例子是基于假设涉及变化的是分立的元件。会论证去除和增加元件，因为这两种情况都会导致学习曲线中的阶梯。

增加

当在生产或安装中，把一个新的元件添加到装配中，分析这个添加的影响时必须考虑几个关键因素。合理的假设是在装配中增加新的组件会导致额外的工时。然而，对这些增加的工时可以做两个一般的假设：

- 增加的组件的学习速率可能与这个单元中的其它部分相同，因为在大多数情况下，元件会类似，而且工作环境（如公司政策、管理方式等）是足够稳定的，因此可以期望学习速率相同。
- 先前对这个单元的学习没有针对添加的元件。对于新的元件，必须发展出一个 T1。

从上面内容，可以认为一个添加应该被处理为与原来的单元有相同斜率或改进速率的一个新的学习曲线。

去除

去除，最简单的情形，是从在产的物品中去掉一个元件。然而，可能需要关注不仅是分立元件的去除带来的变化。例如，去掉那些需求不迫切的工作、实现等。这类变化会对在产的物品的成本产生影响。当然，在这个例子里，成本会减少，因为修改后的单元需要付出更少的努力。换句话说，与去除发生前相比，要做的工作变少了。

案例四：曲线的阶梯函数

背景：

T1、斜率、发生阶梯变化时成本下降或增加的百分比、要预估的份中的单元数是给定的。假设元件在单元 51 处去掉。

数据集合

内容	T1	斜率
带有元件的第 51-75 单元	1000	80%
从 T1 去掉元件	T1 的 10%	80%

计算：

目标：

第 51-75 单元的成本合计使用累计平均理论。

- 步骤一：使用下面的公司计算假设的第一个单元成本T1'：

$$T1' = T1 \cdot (1-P)$$

其中：

P = 第n个单元单元成本减少的百分比

在这个例子中，T1' = 1000*(1-.10) = 900。

- 步骤二：计算第51-75单元的累计数值合计（CTV）：

$$CTV_{51-75} = CTV_{75} - CTV_{50}$$

$$= 900 *(75.678072 - 50.678072)$$

$$= 4042$$

其中：

CTV_{75} = 75个单元的累计数值合计

CTV_{50} = 50个单元的累计数值合计

4.2复杂学习曲线问题的常用提示

- 要做更为复杂的单元曲线的提纲，首先要做一个简单的曲线，然后作为几个单独的曲线问题，逐个解决，加和得到成本合计。
- 通过计算变化前和变化后的成本，来计算包含了斜率变化和阶梯向上或向下函数的份的成本合计，然后把它们加和。这通常需要多次使用上面描述的方法、计算一个或多个用于变化的点之外的成本的其它T1值。如果涉及阶梯函数和斜率变化，需要考虑Y和斜率变化。
- 需要找出单一完整系统有不同单元数量的问题，比如多引擎飞机的引擎，通过把从多个曲线得来的成本元素加和计算成本合计。

4.3生产效率的变化

学习曲线应用程序经常忽视的一个领域是效率变化的影响。正常情况下，在生产计划中使用一个由分析人员给定的生产效率。因此，当分析学习曲线历史记录时，经常会忽视生产效率这个变量。多个研究得出的结论是生产效率是一个成本敏感的变量，应该被考虑进来。政府和行业协会研究学习曲线概念和应用（空中发射武器系统成本工作委员会）的普遍意见是，生产效率确实会影响学习曲线应用。在研究中引用的理论基础确实反应了基本经济学理论。下面总结的这个理论可以作为理解生产效率影响成本缩减的基础。

这个经济学理论建议每个输出有一个平衡，当输出超过这个平衡时，成本会增加，简而言之，因为：

- 新雇用工人的较低劳动生产率；
- 加班带来的额外劳务成本；
- 负担过重的固定设备；
- 不断增加的管理复杂性。

长期运行时，增加的效率会产生一个新的平衡。这是因为：

- 劳动的专业性；
- 大量的已生产的单元吸收了固定成本；
- 管理效率。

如果生产效率下降到低于原来的平衡点，单元成本会增加，简言之，因为：

- 由于裁员或人员去做其它工作而造成的劳工专业性流失；
- 固定设备利用率不足；

- 已经生产的单元数量较少地吸收固定成本;
- 管理效率低。

长期运行时，会在较低的劳动生产率下发展出一个新的平衡点。

不是所有的经济现象都能在学习曲线中体现。表现为固定或半变动成本的限制条件（如装备）导致的效率变化，会极大地影响成本。然而，学习曲线只考虑了变动成本。因此，这个讨论是关于变动成本和它们的影响，这个影响是生产效率变化的结果（新雇用工人的更低的生产力、超时带来的额外劳务成本、劳工专业性的流失）。

在 1976 年，Larry L. Smith 进行了一个研究，是关于在飞机机身生产效率中的变化导致的直接劳动需求的变化。在他的研究中，Smith 开发出了一个累计生产和生产效率计算机程序。这个程序包括了一个生产效率变量，作为两个标准学习曲线变量的补充。Smith 通过把生产效率变量引入基本学习曲线，显著改进了预估。这可以从统计度量中得到证明。

后来使用实际数据的大量研究，都导向了检查生产效率对单元成本的影响。这方面研究的结论已经混合在一起了。任何预估者在面临包含生产效率变化的环境下的预估时，都建议咨询生产效率对成本预估影响的最新研究。

5. 学习曲线评价标准

为了正确评价学习曲线，分析人员必须考虑多个标准。分析人员必须理解在什么条件下使用学习曲线来预估成本，而这要求对学习的原因有很好的理解。分析人员还必须懂得当选择学习曲线参数时要考虑什么因素，以及学习曲线的输出是如何受这些参数影响的。这些概念在本节讨论。

5.1适合使用学习曲线的预估情景

自二十世纪三十年代发布了关于飞机行业的学习曲线的第一篇论文以来，在这个课题上已经写了很多论文。在《*The Learning Curve: Historical Review and Comprehensive Survey*》中，Louis E. Yelle 提供了 90 多篇 1967 年以前发表的参考资料。他最重要的发现是观察到成本在以一个可预测的行为随物品生产数量的增加而减少。基于大量试验研究的学习曲线理论，声明随着用重复工艺生产的单元数量加倍，生产加倍数量的成本会下降一个常量百分比。这个技术最适合用来估计来自劳务和其它因素（如材料加工或管理）的成本减少，这些因素在工艺中是不断重复的。重复的工艺涉及了手工劳动或神经练习，可以从简单到复杂。学习曲线传统上是用来估计制造物品的成本，但是它也能用来估计设备安装成本或任何其它重复工艺的成本。

另一个预估的场景问题是使用学习曲线来预估成本会在什么层面发生。这个问题依赖于预估的全部结构。如果使用了一个顶层参数预估来预估一个开发物品的成本，那么学习曲线可能会用来预估生产成本合计的钱数。另一方面，如果有详细生产数据，那么学习曲线可以用在一个较低的层面，比如预估制造的劳务工时。然后这样的预估会通过应用劳务费率和一般管理费率转为钱数。分析人员必须确定这个预估的性质。本章中依赖的变量样例使用的数据以工时为单位，因为工时只与学习相关。

对于货架物品，使用学习曲线可能不合适，但是可以用来调整数量折扣。这是因为货架物品已经被生产了很多次，因此工艺很标准，学习的影响变得很小。在这种情况下，使用数量折扣是一个更好的方法。数量折扣的讨论会在第六节。

5.2因为学习而造成的成本降低

成本降低的原因是个体的学习和整个组织在重复过程。那么很清楚，只有重复发生的成本受学习影响。非重复发生的成本，如购买工具的成本，不受学习影响。非重复发生成本可能会影响学习效率，但是它们的购置成本不会受对工艺学习的影响。换句话说，如果一个制

造商投资在工艺改进工具上，他可能期望工艺的学习效率发生变化。然而，学习效率不会影响那个初始工具的成本。

按照Rodney Stewart在《*The Cost Estimator’s Reference Manual*》中表达的观点，造成成本降低的因素在学习曲线中通常包括：

- 操作人员学习；
- 改进的方法、工艺、工具、机器和设计以提高生产率；
- 管理上的学习；
- 工程数据除错；
- 生产效率；
- 装配或零件的设计或修改
- 工艺的规范或设计。

通过检查这些因素，很明显，期望操作人员学习而导致工时减少是合理的，改进的方法也是如此。然而，其它成本如材料成本，也能随学习而降低。随着操作人员重复工艺的学习，他们可能找到减少裁切和材料浪费的方法。因为改进工艺而减少人员的劳动强度或增加安全性，可能会令人员周转成本降低。许多因素能归因于学习，而学习可以是个体，也可以是整个组织。这把我们带到了下一个主要的考虑，预估人员必须使用学习曲线——如何选择关键的学习曲线参数。一旦分析人员确定学习曲线是一个建立特定预估情景成本模型的恰当工具，下一步的使用学习曲线主要考虑的是选择适当的参数来建立特定预估模型。这个话题在下面讨论。

5.3数量折扣

本节的讨论迄今位置集中在各种学习曲线技术的应用以及为什么“学习”现象会随着生产单元的增加而降低单元成本。然而当数量增加时，市场中有其它力量也会允许每个单元的成本降低。例如“数量折扣”效应，消费者几乎每天都会接触到的一种价格策略。消费者常常会面对“3毛9一听豆子，1块钱3听”、“买三个轮胎送一个轮胎”。当数量增加时，这些例子里的平均单元成本降低了，而不是因为“学习”。物品是标准的，并且已经大量生产而达到生产工艺的“标准时间”了，没有学习的空间了。在这种情况下，降低成本的主导因素可能变成数量折扣，制造商愿意做这个生意的原因是：

- 竞争
- 规模经济
- 每个单元固定成本降低

数量折扣不只是零售领域的孤立现象，在获得其它物品时也能范县这个现象。例如，电器制造商可能会给一个无线电设备报价2万元，这是购买10台以内的价格；如果购买11台或更多，价格是1万8千。原料提供商报价类似，基于每个定单的数量合计报价。一旦知道了一个制造商或一个销售商的折扣率，可以为几个数量计算单元成本。预估人员应该对历史数据中的数量折扣影响特别敏感。历史数据通常是按生产或采购的时间顺序来做报表和分析的，因此可能会描绘出“学习”现象。

数量折扣曲线的应用不同于学习曲线。使用学习曲线，来自先前的经验会在后续要完成的任务中起到累计效应。结果，单元成本或份购买值是由学习曲线计算出来的，以前单元或份购买的累计效应会起作用。换言之，如果第一份是20而第二份是50，在计算第二份时会考虑生产单元1-20时获得的成本信用。然而，当涉及数量折扣时，信用不会从之前的活动中获得。使用前面的例子，前20个单元的成本会在曲线上从单元1移动到单元20来计算，然后乘以20倍的折扣后的单元成本。对于下一个50个单元份，仍然是从单元1开始沿曲线到

单元 50，然后乘以 50 倍的折扣后的单元成本。

在大多数情况下，只是一个简单的向制造商或销售商的询价，就可以确定是否报价会随数量而有折扣。目录价格更简单明了，它们清楚描述了大量采购时的数量折扣。当数量折扣是影响单元成本减少的主要因素时，应该使用上述方法而不是学习曲线。使用数量折扣理论预估，使用下面的等式：

Equation 18-11 Equation for Projecting Quantity Discounts

$$TC = VC \cdot Q + FC$$

Where:
TC = Total cost of lot of Q units.
VC = Variable cost per unit.
Q = Quantity of lot or purchase.
FC = Total fixed cost for lot or purchase.

5.4 选择学习曲线参数

学习曲线有两个关键参数来决定预估：

1. 学习曲线的斜率（代表了因学习而产生的成本下降）
2. 学习曲线的第一个单元数值

学习曲线参数一：斜率

一个工艺中的期望学习的量是一个分析人员需要考虑的关键。如果分析人员假设 80%的改进率而不是 70%，那么成本上的差异是显著的。假设 80%的学习曲线，如果单元 100 实际成本是 10 万元，那么单元 200 的成本会是 8 万元。对于 70%的学习曲线，单元 200 的成本会是 7 万元。因此斜率的选择必须合理可靠的。有一些学习曲线斜率的来源：实际生产数据、类似工艺的经验以及行业标准学习经验（1985 年版的《电器行业成本预估数据》）。学习曲线存在于各行各业，从软件开发到电子。如果生产商存在同一个工艺的数据，那么分析人员可以提供历史经验斜率，或者分析人员使用实际数据计算出斜率。本节解释如何从实际数据、运用最小方差法计算出斜率。

如果没有同样工艺的实际数据，退而求其次的选择是找到一个类似的工艺。关于如何调整斜率来解决工艺上的差异的工程输入将解决这个问题。例如，如果增加了工艺中的自动化程度，学习曲线会趋向于变得平坦（有更高的斜率值）。原因是机器不会学习——任何与学习有关的机械设计是在这些机器中或在系统的工程成本中。特别是这个成本会典型地体现在理论上的第一个单元成本，这是学习曲线等式中的第二个参数。

最后，一个分析人员会倾向于使用一个行业标准的学习曲线。重复性的电器制造、重复性的机加工操作甚至杂务工作都有行业标准。

一些通用的关于学习曲线的观察包括：

1. 正常的人员利用率与学习曲线斜率关系非常小；
2. 通常，学习曲线斜率会随生产的复杂性增加而增加（准确地说是低于 100%）；
3. 商业的货架物品经常比加工劳动斜率更平坦。

第一个观察被这样的事实所解释：大多数人的学习效率基本一样。第二个观察可以用更复杂的产品会牵涉到更多的学习来解释。第三个观察的原因之一是提供货架物品的商家通常在他们的学习曲线的浅比例，因此，购买者体会不到成本上的大的变化。第三个观察的另外一个原因是货架物品在它们的价格中有很多不受学习曲线影响的部分，因此斜率很平坦。

在选择斜率时，分析人员还需要一个理论上的第一个单元成本（T1），从这个值来开始引申的成本降低。需要注意，只有斜率不足以进行成本预估——还必须有一个初始成本，然后来计算降低的成本。

学习曲线参数 2：理论第一个单元成本（T1）

如同斜率的情况，T1 值可以用多种方法获得。这些方法包括从实际数据获得 T1、从类似或成本预估关系估算出 T1、或者通过使用一个详细的工程设计预算来预估 T1。案例一是如何从实际数据获得 T1 的例子。对于成本预估关系、类似物或者详细预估的具体讨论，请读者参考这些主题的章节。

斜率和 T1 是学习曲线等式的关键参数。两个参数都很敏感——例如，斜率的微小差异都会因数量大而在成本合计中造成很大的差异。

下表描述了斜率从 85%增加两个百分数到 87%，成本合计的变化。相对 85%而言斜率的变化率是 2/85=2.35%。这个表使用累计平均理论，假定的 T1 值是 1.0。

斜率从 85%到 87%的敏感分析结果

生产	生产成本合计		成本增加的百分数
	斜率 85%	斜率 87%	
1	1.0	1.0	0.0
10	5.8	6.3	8.6
100	34	39.6	16.4
1000	198	249.6	26.1
10000	1153.8	1571.6	36.2
100000	6724.7	9895.5	47.2

在使用学习曲线预估重复发生的成本过程中，当 T1 发生了一个给定百分比的变化，很容易看到成本合计按同样的百分比变化。然而，学习曲线的变化则不是这样。对于大的数量，成本合计的百分比变化会比斜率百分比变化大得多。而且，对于给定的斜率变化，成本合计的百分比变化会随原来的斜率和要生产的单元数量而变化。

准确度

使用学习曲线作为预测工具如同任何参数预估工具。度量的准确度，如预估的 R^2 和标准误差，能告诉预测人员对学习曲线能有多大的信心。通常，期望 0.95 或更高的 R^2，而且预估的标准误差越不敏感越好。

5.5劳务金额与工时

当预估劳务成本时，常常是使用基于时间的学习曲线比金额要好，因为金额里包含了通胀或通缩的影响，可能导致效率不遵从同一个学习曲线。如果分析人员选择使用劳务金额作为构建的基础，从历史数据得到改进效率，他应该确信数据已经按经济影响如通胀进行了标准化。

即使在这方面很小心了，预估者应该意识到这样一个事实：大多数审核机构随着时间推移已经对基于工时数据的不同功能成本分类很熟悉了（例如，制造劳务：85%、工程劳务：75%）。如果一个预估是基于劳务金额而不是工时的学习曲线，那么斜率会偏离那些典型的、与明确功能成本分类关联的斜率。随着工作接近熟练，受下降的平均劳务比例的影响，这种

偏离最大的可能性是变得斜率更陡。发生这种平均劳务比例下降是因为开始时会引入高级劳务，随着劳动队伍的建立，会引入初级劳务，因此降低了平均劳务比例。随着劳动队伍稳定成为熟练工作，劳务比例会变得平稳。这个效果加上正常的学习，会在使用金额而不是工时时，令学习曲线变陡。为了避免混淆，估算者应该始终向审核机构证明曲线计算的基础以及历史数据支持所选的斜率。

5.6技术和自动化的影响

随着时间推移，系统和制造工艺已经变得更加复杂。各种进步诸如集成电路和机器人被引入到系统设计和制造环境，影响了学习曲线。然而，这没有在预估过程上改变学习曲线的应用。特别是技术和自动化对“学习”的影响，是使用先进技术和由高度自动化的设备制造的系统演化来的数据库的固有特征。然而，从这类数据计算出来的这些曲线采用本章前面几节所描述的同样方法。

没有简单的经验数据或数学等式允许把曲线直接调整为反应技术和自动化影响。一个直觉的判断是在高度自动化环境中生产的系统会比高度手工生产的同样系统有更平坦的学习曲线斜率。这是因为初始效率的高水平和自动生产设备的因学习而改进的机会减少（因学习而改进在自动设备上会因更高效的装卸和改进了的软件程序而变得很有限）。然而，即使有了这样的直觉知识，影响程度取决于自动化程度。因为每个系统的技术和生产方法是独一无二的，所以对预估者的挑战是从历史记录数据库中选择与要预估的系统尽可能一致的技术和制造的那些系统。这可能会需要分析多个系统，以及所分析系统的方方面面是如何影响的组合度量的。最后的结果会是一个对所用技术和制造环境的判断：对这个系统而言，什么样的T1和斜率调整是合适的。

下面是对会出现的、与学习曲线有关的典型预估问题的描述。

5.7基于学习曲线的数据的可靠性

在第六章的6.3和6.4节“为价格分析收集和评估数据”中包括了一个影响数据可靠性的因素的讨论。这个讨论能够扩展到学习曲线。相应地，在应用学习曲线模型前，分析人员应该评估数据的可靠性。

对数据的常用考虑包括：用时、质量以及历史记录数据。要想有用，数据必须系统地、及时地进行采集。此外，数据必须质量可靠：准确和贴切、完整、通用。最后，历史数据必须根据行业、通胀等进行调整。

6．总结

学习曲线是有用的预估和分析工具。基础概念和评估规则的知识，令分析人员能对评估学习曲线提供预估是否有效进行判断。除了单元和累计平均理论，分析人员应该考虑特殊情况，如区分学习曲线和生产效率变化。当评审要使用的学习曲线时，分析人员应该检查预估情景、成本降低的原因、曲线参数的选择、技术的影响以及曲线所基于的数据的可信度。

一元线性回归算法说明

从上面对学习曲线理论的说明中，我们知道可以采用一元线性回归算法从历史数据中获得可信的学习曲线数学公式。

在把一元线性回归应用于学习曲线时，有两个技巧：

1. 最小方差应该大于等于 0.95，因此在最小方差小于 0.95 时，应剔除偏差最大的数据，重新计算，直至最小方差大于等于 0.95 为止；
2. 对于没有历史数据或历史数据数量不够的新人，在预估其工作量时，斜率可以采用同类工作的平均斜率。

功能列表

本组件应实现以下功能：

1. 根据给定的最小方差、历史数据、新的数量按照 Wright 法计算出预估值。
2. 根据给定的最小方差、历史数据、新的数量按照 Crawford 法计算出预估值。
3. 根据给定的最小方差、历史数据按照 Wright 法计算出斜率和 T1 值。
4. 根据给定的最小方差、历史数据按照 Crawford 法计算出斜率和 T1 值。

接口及优化

本组件为其它模块（组件）提供静态调用方法，没有需要存储到数据库的数据。

接口类是 cn.langhua.ofbiz.learningcurve.I_OFBizLearningCurve.java，请开发人员实现该接口类中的方法。

－－　本文档结束　－－