

关于强化风险管理和金融机构经营的系列报告

以内部评级制度为基础强化信用风险管理

2005年7月

日本银行金融机构局

联系人 日本银行金融机构局

福本 智之 (Tomoyuki Fukumoto)

电话 +81-3-3277-3019

电子邮件 tomoyuki.fukumoto@boj.or.jp

## 目 录

1. 序文.....	3
2. 内部评级制度概要 .....	4
(1) 内部评级制度的定义 .....	4
(2) 内部评级制度的意义.....	4
3. 内部评级制度体系.....	7
(1) 债务人评级和项目评级.....	7
(2) 实施评级的对象.....	9
(3) 评级的划分.....	10
(4) 评级的评估期限——实施评级和景气循环的关系.....	11
4. 评级实施程序.....	13
(1) 评级的实施和修正程序.....	13
(2) 债务人评级中的定量评估和定性评估.....	14
5. 评级模型.....	16
(1) 评级模型概要.....	16
(2) 评级模型的验证.....	17
(3) 评级模型的修正.....	23
6. 风险要素推算.....	24
(1) 风险要素推算的基本程序.....	24
(2) 风险要素的查证.....	27
7. 内部评级制度的灵活运用 .....	30
8. 内部评级制度的验证.....	32
9. 信用风险的计量化.....	36
(1) 预期损失(EL)和非预期损失(UI).....	36
(2) 风险计量.....	38
(3) 压力测试.....	39
(附件 1) 信用风险管理体制和强化信用风险管理设想.....	41
(附件 2) 为实施验证的统计数量事例.....	42
(附件 3) 违约时损失比率(LGD)推算事例.....	44

## 1. 序文

2001年10月，日本银行发布了主题为『灵活运用信用评级，完善信用风险管理体制』的报告，本报告为该报告的延续。在此期间，金融机构的信用风险轮廓发生了很大变化，其管理方法也取得了各种各样的进展。同时，随着『巴塞尔II<sup>1</sup>』的正式出台，现在我国正在为适应该规则准备着国内法规。以下内容将根据这些情况的变化，就以内部评级制度（信用评级制度）为基础的信用风险管理现状及课题的主要论点进行介绍。今后，日本银行的考查和监管将以这些论点为核心，与各金融机构进行深入探讨，促进信用风险管理的日益完善和强化。

本报告由以下内容构成。首先，第2章介绍作为强化信用风险管理基本工具的内部评级制度的概要。接下来的各章节介绍以该制度为基础的信用风险管理现状和课题，按照具体的论点展开说明。主要论点包括：内部评级制度体系（第3章）、评级实施程序（第4章）、评级模型（第5章）、风险要素推算（第6章）、内部评级制度的活用（第7章）、内部评级制度的考证（第8章）。在此基础上，第9章介绍内部评级制度的完善过程，整理信用风险计量的思路。再适当通过专题论述的方式，介绍对于经常遇到的疑问的想法。有些方法现在并没有正式确立、普及起来，但是今后可能会被重点采用的论点，所以在这里也一并介绍。

本报告所设定的信用风险管理体制和信用风险管理强化的整体思路，是通过附件1体现的，敬请参考。

---

<sup>1</sup> 巴塞尔银行监督委员会于2004年6月公布了『自有资本测定和关于基准的国际统一化：改订框架』（International Convergence of Capital Measurement and Capital Standards: A Revised Framework）。

## 2. 内部评级制度的概要

### (1) 内部评级制度的定义

内部评级制度是指，金融机构在做信贷业务时，以控制信贷风险、强化管理为目的，根据的债务人及交易对象对风险进行分类管理的系统。

迄今为止，许多金融机构通过对个别债务人管理的方式，来解决自身承担的信用风险。对于授信判断，也将『贷款』或『不贷款』作为了重点。而且，即使债务人破产了，也大多可以通过房地产抵押全额回收贷款。

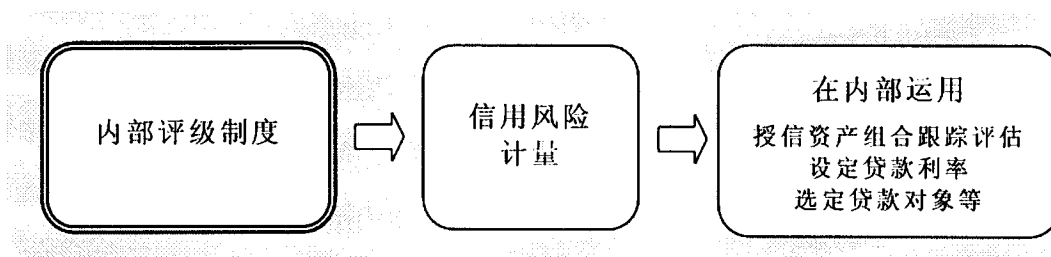
但是，在所谓的『泡沫经济破灭』以后，破产债务人日益增多，抵押的价值也日益降低。在这种背景下，许多金融机构的信用成本开始增大，如何控制信贷风险成为了重要经营课题。随之而来的是要求有高精度度的信用风险管理手法，以内部评级制度为基础的信用风险管理和统计方法开始普及起来。尤其在 1998 年引入自身查定制度以后，作为判定债务人分类的一种机制，内部评级制度开始被采用。包括中小金融机构在内，将这种机制灵活运用于信用风险管理活动的机构日益增多。

### (2) 内部评级制度的意义

通过完善内部评级制度，金融机构可以完成哪些事情呢？首先，对于个别授信的判断和中间管理等可以实现高效化。其次，在确保债务人信用级别和授信交易质量方面，可以按照统一的尺度进行把握。再次，通过分析内部评级、授信余额及项目数目的推移，可以对授信资产组合整体质量进行评估。通过的评级违约概率的推算，可以实现信用风险的计量管理。这种对信用风险的统一、全面、客观的实施和把握，可以成为确保稳健经营的基础要素。

此外，完善内部评级制度的好处，并不仅仅体现在管理层面的改善上。在经营战略层面，根据的信用级别，判断信贷利率如何确定，判断通过扩大什么样的评级区域的信贷规模才能从风险和回报的角度得到更好的收益效果。通过类似的判断，可以获得更加客观的材料（参照图表 1）。

(图表 1) 信用风险管理强化程序



通过这种方式,我们可以看到内部评级制度可以成为强化信用风险管理的前提。所以,几乎所有的金融机构都需要进一步完善该制度。但是,各金融机构所处的环境,按其所处的环境,在要求建立的内部评级制度设计上实际也各不相同。比如,对于中小金融机构,同大的金融机构比较,简单的内部评级制度显然更为适合。内部评级制度结构(评级的划分数、各等级评级的定义、评级实施方法等),并非只有一个唯一的绝对的答案。我们希望各金融机构根据自身的授信资产特征、业务内容、评级制度的主要应用目的来进行判断,从而建立符合自身特点的评级制度。当然,随着外部环境的变化,内部评级制度本身也需要不断地灵活地予以修订和完善。

如上所述,内部评级制度有各方面的内容,本报告,就在内部评级制度的构筑和运用时的重要论点,如下整理我们的思路。

- 内部评级制度体系
  - 构筑内部评级制度的主要框架时,哪些要点必须留意并做出判断?
- 评级实施程序和评级模型
  - 债务人、授信项目的有关信息如何同评级相结合?
- 风险要素的推算和信用风险的计量
  - 如何灵活运用实施的评级结果,推算风险要素和风险的大小?
- 内部评级制度的灵活运用和验证
  - 如何构筑能够灵活运用并且能够经常改善的机制?

### 【专题 1】 建立内部评级制度的必要性是否因金融机构的规模不同而异？

关于建立内部评级制度，『所面对的债务人数量很少的中小金融机构难道不是没有必要吗？』，有这样的意见。出现这种意见的背后，主要是基于『债务人数量较少时，通过统计手法管理信用风险没有效果』这样的认识，以及『内部评级制度的构筑、运行需要巨大的成本投入』等担忧的一种体现。

但是，即使规模较小的金融机构，一般情况下债务人数量也会达到相应的规模。而且，在各个债务人的信用程度有很大差距的时候，信用质量相同的债务人分类至同一个信用级别内进行管理，会对授信工作的推进带来好处。因为通过明确确认债务人的评级，可以客观相对比较地来评估其信用程度。尤其是对于信用级别低、或者信用级别下降特别明显的债务人，可以充分认识到其风险的同时，还可以及早研究讨论对这类个别债务人的应对措施。比如，强化对这些债务人的跟踪监管，可以根据情况需要采取有张有弛的风险管理体制。而且，通过对所有债务人按照信用级别不同进行分类管理，自己还可以整体分析、把握自身的授信资产组合状况。也就是说，是否授信集中于某些特定的评级领域；各评级等级间的变化程度等等，通过这些观察可以进一步分析授信资产组合的质量及变化情况<sup>2</sup>。

当然在债务人数量很少的时候，有一种观点认为，『即使不依赖内部评级制度，通过对个别债务人的管理，经营者及信贷负责人可以掌控住经营出现危险的债务人和授信资产组合整体的现状』。但是在这种情况下，关于债务人及授信资产组合整体的信息，只是集中在组织内的一部分人员手中，从信息共享、保持经营的连贯性等观点出发，这种做法是不可取的。或者对利害关系的所有者、监管当局履行说明责任时，存在不能及时体现高效等问题，这些都需要我们统一进行考虑。

另外，因为构筑内部评级制度及其运行需要一定的成本，所以，引入什么样的制度，需要对其成本和效果进行彻底认真的分析。比如，如果认定早期即引入正式的内部评级制度并不合算，则可以先引入简易制度，之后再根据环境变化和经营战略的调整等，分阶段地进行升级改造。

<sup>2</sup> 再进一步讲，如果持有准确度的内部评级制度，则可以推算出显示授信资产组合特征的特征参数（比如：按不同信用级别算出的违约概率）。

### 3. 内部评级制度体系

#### (1) 债务人评级和项目评级

内部评级制度，是依据信用级别的高低对债务人进行分类的制度。一般包括：①对债务人自身信用进行评估的『债务人评级』、②对个别授信项目进行评估的『项目评级』2个种类。前者是对交易对象债务人的信用风险高低、即根据债务人严格履行合约还是违约（不履行债务的可能性）的程度进行分类（图表2）。这是评级体系当中最为标准的内容，许多金融机构已经开始采用。另一方面，后者是指着眼于个别交易风险而进行的分类，以债务人的信用质量，加上抵押、担保等保全手段及授信期限的设定等进行评估。所以，即使是同一个债务人的授信，根据具体项目的不同评级也会有差异。

（图表2）债务人评级的等级划分（例）

评级划分	定义	债务人划分
1	财务状况健全，切实履行债务的程度最高。	正常
2	财务状况良好，切实履行债务的程度高。但是，业务环境发生大幅变化时，债务履行的确切性有可能降低。	
3	财务状况基本良好，从当前情况看切实履行债务不存在问题。但是当业务环境发生变化时，债务履行的确切性有可能降低。	
4	财务状况基本良好，从当前看切实履行债务没有问题。但是当业务环境发生变化时，债务履行的确切性降低的可能性较大。	
5	虽然可以确认履行债务的确切性，但是当业务环境发生变化时，可以看到损耗履约能力的要素。	
6	切实履行债务的前景不够充分，如果业务环境发生变化，则履约能力将有可能受损。需要注意业务发展状况。	关注
7	业务状况、财务内容存在问题，影响债务履行状况的障碍大。	
8	业务状况、财务内容有重大问题，债务履行状况发生问题或者处于比较接近的状态。	
9	处于经营困难状态，经营改善计划等进展不顺利，今后陷入经营破产的可能性大。	可疑
10	处于严重的经营困难状态，甚至实质上已经处于破产状态，或者在法律上、形式上破产的事实已经发生。	实质破产，破产

判断评级时，如果是债务人评级，则债务人违约的可能性是最基本的分类标准，一般使用违约概率（PD）这一指标。另一方面，如果是项目评级，则有必要测试个别交易的指标，它是由预期损失比率（EL 比率）[=违约概率（PD）× 违约时损

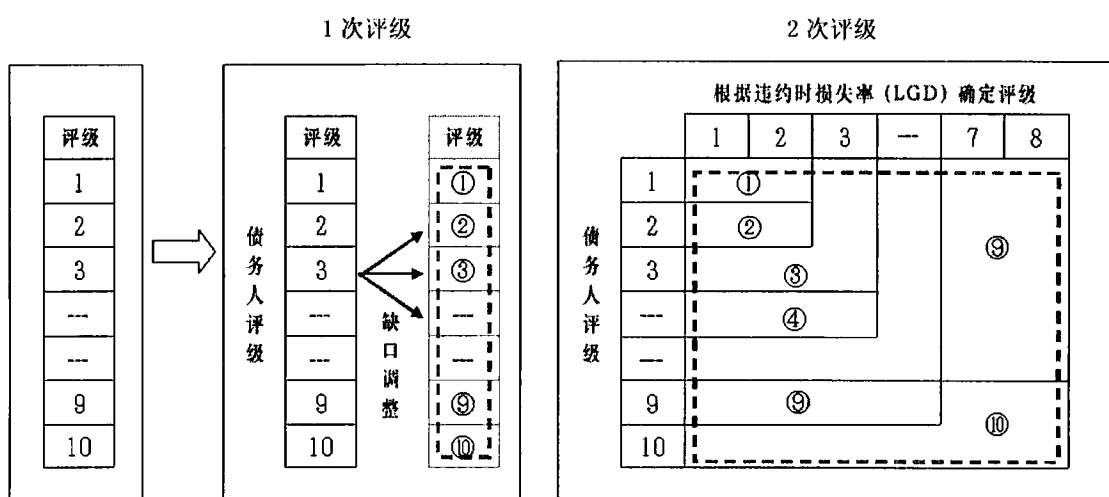
失率（LGD）]来代表的<sup>3</sup>。

项目评级是比债务人评级更为精致化的内容,包括 1 次评级和 2 次评级 2 种(图 表 3)。1 次评级是指以债务人评级为出发点,考虑授信特性(保全状况等),通过评级的上下缺口调整确定项目评级的方法。而 2 次评级是指同债务人评级无关,独立评估项目特性的评级体系,比如根据违约时的损失率(LGD)设定级别,再同债务人评级结果相结合,从而确定项目评级的方法。

(图表 3) 项目评级设想

【债务人评级】

【项目评级】 ① ⑩是项目评级的符号



在我国,面向企业授信为中心,债务人评级被广泛引入。项目评级的对象往往只限于被认为以项目为准进行管理比较适合的特殊债权,比如项目融资、房地产信贷,证券化项目等。

对企业授信的项目评级没有普及的原因之一,在于无限抵押、无限担保这种我国存在的特有的制度。存在无限抵押、无限担保的情况下,抵押、担保无法和个别项目联系在一起,因此对于个别项目单独判定其信用程度是比较困难的。

<sup>3</sup> PD 是 Probability of Default、LGD 是 Loss Given Default、EL 是 Expected Loss 的简称。这些推算方法请参照第 6 章。



## (2) 实施评级的对象

原则上所有的债务人、项目都应该成为评估的对象。由于金融机构信用风险的评估主要集中体现在评级上，从风险管理的精致化、效率化的观点出发，对于所有的信用风险全部进行评估是最为理想的。但是根据交易的规模和特性不同，对所有的债务人、项目均进行评级，从成本的观点出发并非都是合理的。那时，可以不进行评级等例外的运用模式将被认可，当然需要事先确定这类债务人和项目可以不进行评级的基准。除此之外，所有大额交易的对象、项目都必须成为评级的对象。

所说的零售领域的授信，有小额分散的个人贷款及无抵押的固定形式的商业贷款。对于这一类贷款，与企业的授信方式，可以对每一个债务人不进行单独评级，而是分成几个领域进行评估。统一授信的类别，比如按照不同的商品（房地产信贷、信用卡贷款等）或者根据债务人属性的不同（职业、年龄等）等分成几个领域<sup>4</sup>，按照每一个领域进行信用管理的方法（图表 4）。类似这种小额分散式的交易，与其重视债务人每个人的个别特性，更应该重视商品形态、债务人的属性等每个领域的统一特性（比如个别领域的平均违约概率），这样可以更为高效地执行信用风险管理。正是基于这一点，各领域内的风险程度需要具备等质性。完成上述确认时，比如收集各细分领域的风险要素（违约概率等）的实际数值，对其稳定性进行评估，这是最普遍的操作方式。

（图表 4）各个统一授信领域的细分状况（例）

按商品区分	债务人属性	拖欠状况
住房信贷	A 类	无拖欠
		拖欠 30 日以上
	B 类	无拖欠
		拖欠 30 日以上
信用卡贷款	A 类	无拖欠
		拖欠 30 日以上

<sup>4</sup> 此外，可以根据授信执行时期、经过的年数，拖欠天数等进行领域的细分。

### (3) 评级的划分

为了正确掌控信用风险，评级的等级设定是否合适也是非常重要的一点。评级区分的数目、各等级的信用风险必须保持等质性<sup>5</sup>。尤其是在债务人众多、而且债务人的信用情况存在很大差异的情况下，评级的等级划分最好也要多一些。除此之外，即使评级划分的等级很多，但是，如果结果是债务人过于集中于某一个评级区间时，从正确把握信用风险的观点来看也存在问题。所以，在评级划分设定时，要想办法避免这种过度集中于一个评级区间的事情发生。另一方面，在债务人数量较少的时候，如果增加评级划分则可能导致各评级的债务人数目都很少，结果是可能对风险要素的推算精确度带来不利影响。如果出现这种情况，可以考虑在一评级区内信用风险等质性不会带来破坏的条件下，将同相邻评级进行统一合并，或者运用外部的样本数据等补充手段。

同时，即使内部评级制度已经构筑起来，授信资产组合没有特定的评级区间偏颇，分布非常均衡，但是随着时间的推移，由于景气循环的影响，有可能发生整体向一个方向倾斜导致某一评级区间过度集中的情况。出现这种情况时，在重视评级的稳定性的同时，必须彻底分析上述变化持续性，在此基础上探讨适当的评级基准、各个等级划分的调整。

#### 【专题 2】 最合适的评级结构之思路

在确定评级划分方法时，并没有特别明确的基准。不过，比如债务人评级时，确定 1 个评级当中，①债务人违约概率幅度、②债务人数目的多寡是重要的判定要素。

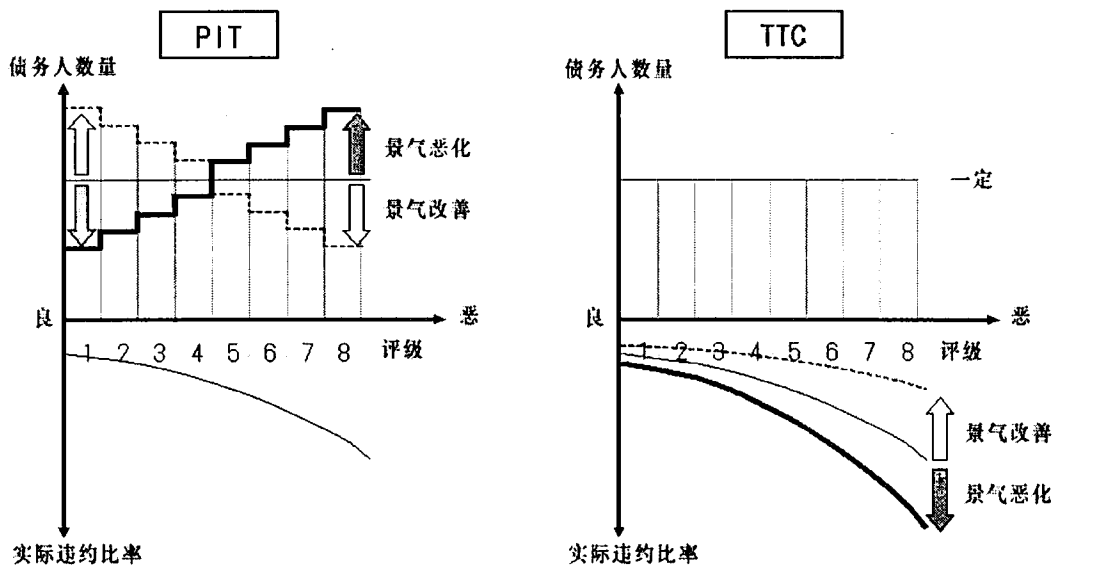
一般来讲上述 2 点属于权衡的关系。也就是说，评级划分区间增多时，1 个评级债务人的违约概率幅度就会变窄，虽然等质性升高，但是该评级所包含的债务人数目减少，从而会导致违约概率的推算误差加大。

如果着眼于这一点，为了实现某个评级划分的『债务人违约概率的偏差』和『债务人违约概率的推算误差』的合算达到最小化，可以考虑确定评级数量及评级划分的方法。通过这种方式确定的评级结构，由于各个评级违约概率的推算值和各个债务人的违约概率之间的差距的总和实现了最小化，所以从这个意义上讲这是最合适的评级结构。

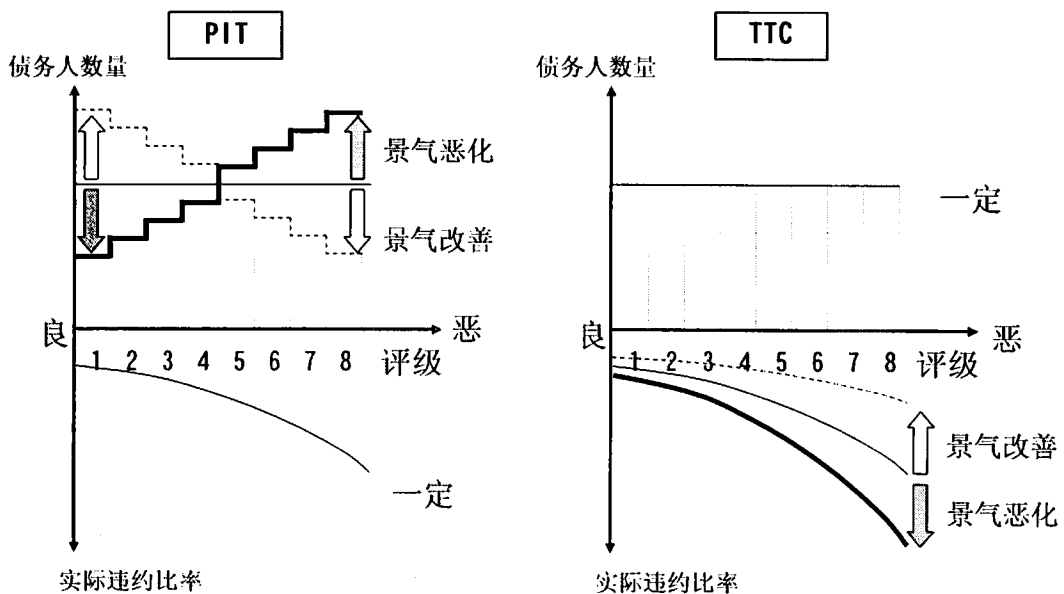
<sup>5</sup> 在巴塞尔 II 的内部评级方法中，要求非违约评级在 7 个阶段以上，违约评级在 1 个阶段以上设定。

(4) 评级的评估期限 —— 实施评级和景气循环的关系

对债务人进行评级的时候，在什么样的一个期限内对债务人的信用水平进行评估也是一个重要的论点。比如，只是对最近的债务人状况进行关注呢？还是根据景气变动的影响从中长期角度对债务人状况进行评估呢？所设定的期限不同评估结果也有可能差异巨大。按照评级制度的种类和景气循环的关系进行分类的话，大体可以分成①Point-in-time 法（以下 PIT）、②Through-the-cycle 法（以下 TTC）这两类。PIT 是不涉及到景气循环局面，根据企业最近期状况进行评级的方式。而 TTC 则是不考虑最近期的企业状况，根据包括景气循环一周期的长期时间内的最恶劣的状况为基础进行评级的方式。



(图表 5) PIT 和 TTC



一般来讲评级公司据说都是从长期视角出发进行评级的，从某种程度讲采用了 TTC 评级的方式。另一方面，从金融机构内部评级的运用状况看，明确地意识到底是 PIT 还是 TTC 开展评级的金融机构很少。他们对债务人几年后（比如 3 年到 5 年）的信用状况进行评估，即采用了所说的 PIT 及 TTC 的中间评估方式。

如何判断 PIT 及 TTC（或者其中间）哪一个更为合适的问题，鉴于 PIT 及 TTC 的差别在于分析债务人信用程度的期限长短上，依靠于金融机构承担债务人的风险期限的长短。如果执行的贷款半数以上是长期信贷的金融机构，则考虑在较长期限内根据信用情况的变化，进行评级是比较合理的。当然，彻底分析并确定债务人未来的信用变化状况并不容易。特别是像 TTC 那样，对所有债务人进行评级时，包括景气循环在内需要分析其长期的变化趋势，这一点从技术上是讲是非常困难的。所以，评级仍然依据最近期的信用状况实施的同时，假设景气出现恶化演练，从而彻底看透风险的严重程度是非常重要的。这种做法可以作为评级制度类型，采用 PIT 的同时，在运用层面上加入 TTC 的要素。

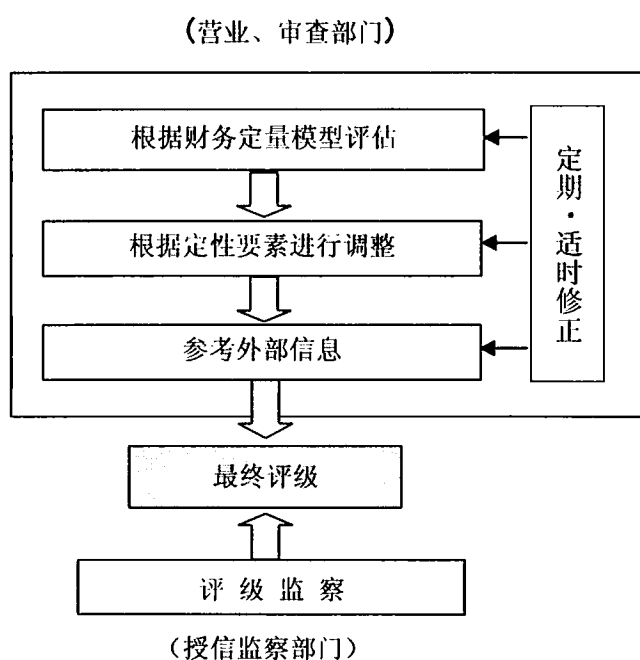
与选择 PIT 还是 TTC 这一问题无关，作为金融机构自身来讲，采用的内部评级制度同景气循环存在什么样的关系？即准确把握是 PIT 还是 TTC，抑或是中间型的问题是非常重要的。为什么这样说，因为对评级制度的精确度以实际违约率为基础进行评估时，是 PIT（不同的评级实际违约率稳定）还是 TTC（针对不同的评级，实际违约率会随着景气变动而变化），其方法会不同的。

#### 4. 评级实施程序

##### (1) 评级的实施和修正程序

实施评级需要按照事先确定的手续，一般由营业部门及审查部门进行操作。届时，通常依据财务定量模型进行评估，根据定性要素（如产业动向等）加以调整，并参考外部信息（外部评级、股价等）<sup>6</sup>这样的流程<sup>7</sup>来操作（图表6）。在此基础上，授信监察部门验证营业部门及审查部门实施的内部评级的稳妥性和适当性。

（图表6）评级实施程序事例



评级并非只是在交易开始时实施，根据债务人的信用情况变化需要适时修正，而且必须经常保证适当准确的评级<sup>8</sup>。所以，在决算期实施的定期修正基础上，判定大额交易对象出现破产迹象等债务人信用发生重大变化时，必须进行适当的修正。

上述的评级实施程序及评级体系，需要以规程或指南等方式明确下来，取得经

<sup>6</sup> 外部信息作为内部评估时的信息，即使不被使用，在进行评级验证时也希望能够被灵活运用。

<sup>7</sup> 内部评级制度不仅作为强化信用风险的一个环节非常必要，而且是准确的自身核定、适当核销和拨备的基础。所以，内部评级制度和自身核定的债务人分类，必须要达成统一。

<sup>8</sup> 假设实施评级采用的是PIT方式，则随着景气循环，信用情况的变化反映在评级上。所以，需要一定频率的修正。

营者的批准后，将其内容通告各有关部门。同时，有关制度设计的基础资料、文件等，必须妥善保管。

## (2) 债务人评级中的定量评估和定性评估

对于债务人评级，一般需要将依据定量信息（债务人的财务指标等）进行的评估，和依据定性信息（产业动向及母公司的支援状况等）进行的评估相结合，进行综合判断。关于定量评估，多利用和企业违约迹象高度相关的财务指标，使用财务定量模型进行分析（图表 7）。为了提高评级的精确度，分析时使用的财务信息，在公开信息的基础上，利用反映财务实际状况的不良资产、非正式损失等信息的情况也很多<sup>9</sup>。

仅仅依据定量信息进行的评估，如果判定没有准确掌握债务人的信用状况，则有必要依据定性信息加以修正（图表 8）。至于修正方法，可以定性要素进行评点，在定量要素的评点处通过加减进行计算，或者以定量要素为基础，采用对评级适当进行升、降级的方法。

（图表 7）决定债务人评级的定量要素事例

规模指标	自有资本额、纯资产额
安全指标	流动比率、资本充足率、经常收支比率
收益指标	净资本经常利润率、销售额营业利润率、带息负债还款年数、利息偿付比率
其他指标	增收比率、利润增加比率

（图表 8）决定债务人评级的定性要素事例

产业特性	增长性、市场行情变动的大小、市场进入壁垒
企业特性	同母公司或资本合作方的依存关系、经营者能力、是否有外部监查

<sup>9</sup> 实际财务信息是指，同表面财务信息比较精确度更高的信息。但是，从信息的客观角度及收集数据的成本问题考虑，财务信息的灵活使用方面，应该充分考虑两者的长处和短处。

对于项目评级，除债务人评级中使用的信息以外，关于该项目的定量和定性信息也是需要的。比如，对通常的企业授信实施项目评级时，是否有抵押、担保，其具体内容，优先、劣后的顺序，授信期限等信息。正像债务人评级当中，必要的信息对违约概率（PD）带来影响一样，项目评级时的必要信息，也同项目违约时的损失比率（LGD）密切相关。

还有，房地产信贷、项目融资等，与通常的企业贷款不同，对债务人不进行评级，大多只对项目本身进行评级。这种情况时，使用图表 9 中列举的定量要素和定性要素对项目的质量进行评估。

（图表 9）决定项目评级的要素事例

	房地产信贷	项目融资
定量要素	授信期限、LTV、DSCR 等 <sup>10</sup>	授信期限、DSCR 等
定性要素	房地产特征（占地条件等）、现金流计划的妥当性、业主风险等	项目风险（业主风险、完工风险、施工风险等）、转移风险等

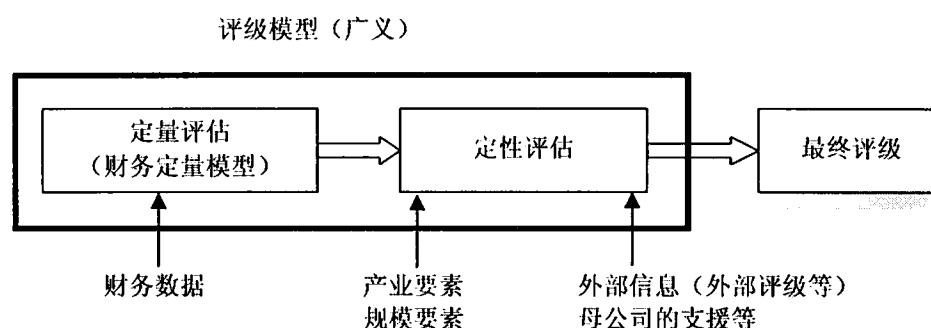
在进行定性评估时，需要事先确定评估基准。一般来讲定性要素，确保判断的客观性和统一性非常困难，容易出现随意性。所以，尽可能将评估基准具体化，详细的文字表述是非常重要的。此外，需要通过内部培训将负责评级人员的理解和认识统一起来。

<sup>10</sup> LTV (Loan to Value) = (授信余额) ÷ (房地产评估额度)  
DSCR (Debt Service Coverage Ratio)  
= (相对应资产运用中获得的当年现金流) ÷ (对债权人支付的本息金额)

## 5. 评级模型

将实施评级的定量评估和定性评估方法予以定型则成为评级模型（图表 10）。评级模型是使用和企业信用有关的信息，来进行客观评级的工具，在内部评级制度中发挥核心的作用。通过引入评级模型，营业部门及审查部门的评级负责人员，不仅可以高效率地进行工作，还可以减少评级负责人之间的评估差异，从而达到进一步提高组织整体实施的信用风险评级的稳定性、客观性的效果。评级模型这一称呼，既包括单指的统计模型，也包括广义泛指评级方法整体。

（图表 10）评级模型概要



### （1）评级模型概要

具体的评级模型，包括各种利用企业财务信息进行信用评估的财务定量模型<sup>11</sup>，得分形式的模型等。而且，财务信息的利用方法，也包括各种数据直接计算方式、按照不同的财务科目设定评点的基础上，将各个评点加入计算方式中。

此外，按照不同的行业、规模、授信级别，根据其属性使用各自的模型是一种方法，重视模型的统一性，使用单一的模型完成所有的评级是另外一种方式。前者由于模型较多，所以采用 1 个模型进行评级时，可以使用的推算样本数据就会减少，结果容易带来模型的可信度问题。另一方面，后者对不同的行业和规模的企业，通过一个单一的模型进行评估，不可能完全捕捉到应有的信息，有可能导致其精确度

<sup>11</sup> 财务定量模型包括判别模型、逻辑回归模型、使用神经网络手法的模型。



不佳。鉴此，将上述 2 者的一长一短认真研究，选择确定对各家金融机构最为合适的方法是很重要的。

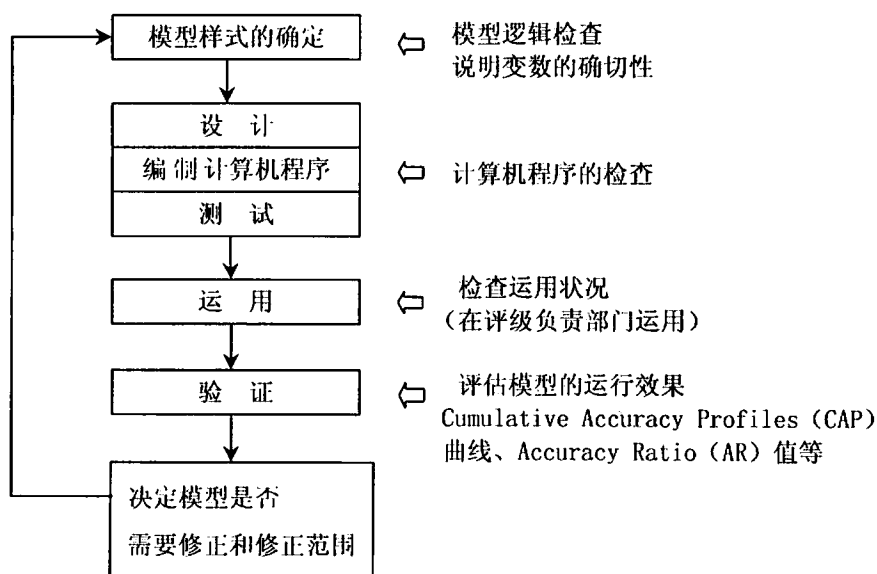
## (2) 评级模型的验证

评级模型对于内部评级制度来讲是非常重要的工具，必须努力提高其精确度。为了实现这一目标，在建立模型时的事前验证和模型运用之后的事后验证，这两项工作必须同时进行（图表 11）。

建立模型时的验证，主要是指模型的逻辑及建立时使用的数据是否确切的确认。当然，使用财务定量模型时，必须验证作为输入变数（说明变数）的财务指标的妥当性。另一方面，模型在运用以后的验证，以债务人的违约发生状况为基础，对模型运行状况进行确认。比如确认评级模型得出的评估（得分）能否准确判断实际发生违约的债务人及没有违约的债务人。

评级模型的验证，并非只是依存 1 个方法，希望能从多个层面进行评估。尤其在信用风险评估领域，关于违约的数据很少，为确立模型所必须的数据是并不充分的。带来的结果就是，由于统计方式的验证一般不容易开展，所以必须在定性方面认真核实，同时，持续不断地完善验证工作<sup>12</sup>。

(图表 11) 评级模型的建立和验证程序



<sup>12</sup> 评级模型中，有复数使用财务定量模型、发挥定性调整作用的几个辅助系统。在这种情况下，各个模型、辅助系统是否在准确正常运转也需要确认。

下面，通过一些具体事例介绍一下在建立模型时，事先验证的方法及实际上使用了违约企业数据后，对模型精确度进行验证的方法。

### ① 模型建立时的事先验证

#### （逻辑的确认）

在判定模型的逻辑是否合适的同时，还需要验证按照这一逻辑建立了该模型。监察部门作为第三者进行验证时，需要确认计算机程序是否按照逻辑的样式编制。此外，有待利用试验数据（假想数据）来进行核实。

#### （验证输入的信息）

对于在评级模型中使用的财务指标是否妥当需要进行确认。具体来讲，首先确认被使用的财务指标的妥当性问题，和企业的信用情况完全没有关系的数据应该排除掉。评级模型如果是由回归模型（逻辑回归模型）建立的，需要对说明系数的正、负是否和逻辑吻合进行确认。进行这些定性确认之外，需要使用各种统计数量（t 值、F 值、AIC 等<sup>13</sup>）定量验证，这也很重要。因为作为说明变数，选择的是相互之间有很强关联的数据，所以确认模型是否存在不够稳定的问题（多重共线性）非常重要。

此外，推算模型时使用的企业财务信息中，如果含有异常数值（从整体的平均看极端背离的数值），有可能影响模型的精确度。这时，需要适当处理并删除这些异常数值。

### ② 使用违约企业数据的事后验证

#### （逆向测试）

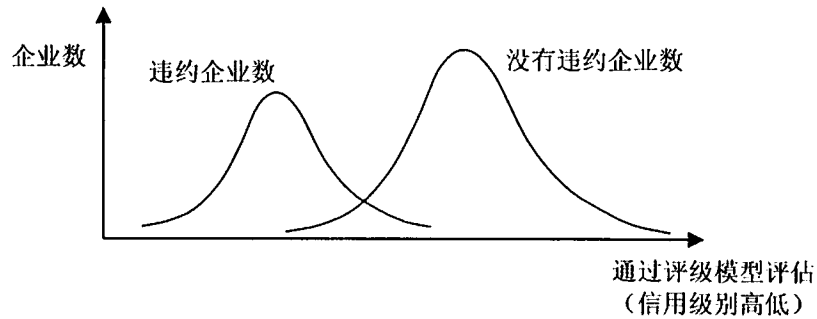
逆向测试是指模型的推算结果和实际情况进行比较，事后确认模型精确度的方法。精确度低的模型，经常会将违约企业判定为信用高的企业，相反，将没有违约的企业判定为信用低的企业。

客观评估模型的精确度，可以考虑按照模型评估信用度（得分）的顺序，排定对象企业，从而调查实际违约企业和没有违约企业的分布情况（图表 12）。如果是判别违约企业能力高的评级模型，由于低级别的企业群体违约倾向高，所以体现在分布图上，违约企业和没有违约企业分布中重叠的部分一定变小。这除了可以通过图解确认以外，从企业的分布当中算出统计数量，客观地开展模型的精确度评估很重要。届时，被普遍使用的定量指标（Accuracy Ratio）事例将在附件 2 中进行整理。

---

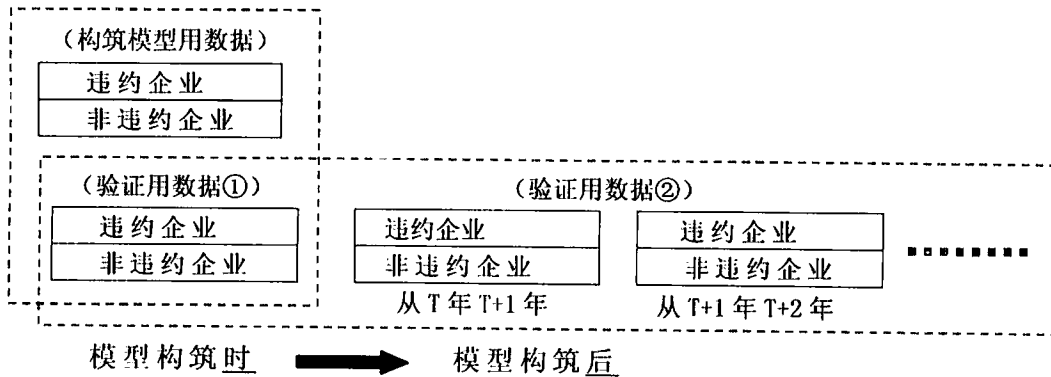
<sup>13</sup> t 值、F 值是使用回归模型评价说明变数有意性的指标。此外，AIC（赤池信息量基准：Akaike's Information Criterion）是指，评价复数说明变数组合情况的指标（信息量基准）。

(图表 12) 通过评级模型评估的违约企业和没有违约企业



**【专题 3】验证评级模型的精确度使用定量指标时的注意事项**

验证模型的精确度，使用 Accuracy Ratio (以下称 AR 值) 等定量指标时，可以使用构筑模型用的数据进行评估，也可以使用验证数据进行评估，必须注意两者的意义是不同的。在使用构筑模型用的数据的时候，即使判定精确度非常高，但是，并不意味着其它的验证数据也一定会得出相同的结果。从验证模型的判别能力的观点看，除准备构筑模型用的数据外，同时准备验证用数据，通过两者的样品计算出 AR 值进行验证。同时，通过不同时期的数据确认 AR 值的稳定性也很重要。即使一个时期的数据可以判断运行态势良好，但是另一个时期的数据如果得出不佳评估结果时，需要对模型进行修正。



### ③ 使用评级变化、违约率进行事后验证

事后验证评级模型精确度的方法，除了前述的使用违约数据的方法之外，使用评级变化、不同评级的实际违约率的方法也值得考虑。比如，被判定为优良、违约概率很低的评级的债务人，在下一期时极度恶化，或者发生违约的情况很多时，表明该评级模型的精确度低。具体可以通过评级行列变化及不同评级的违约概率顺序进行确认（图表 13）。

评级变化行列是指在一段时期内，表明授信资产组合的评级变化状况。对于行列变化，首先检查评级急剧变化部分（异常值），特别指定相对应的债务人。该债务人的评级，受什么要素影响大幅变化进行详细分析，还可以对定量模型、定性调整是否稳妥，甚至模型整体运行进行评估。

还有在评级变化行列中，『从一个评级区向相近的评级区变化，这种情况发生的概率大』，『注意到向某一评级区变化时，则由相近的评级区向该评级区变化的概率大』，以上两种变化情况作为最通常的变化趋势可以观察到。在实际观察评级的变化状况时，确认上述变化规律，如果出现不规则变化则需要探究其原因，这也是一个可以考虑的验证手段。此外，使用违约概率，判定其实际数值是否同评级划分相吻合，也就是在评级划分区内，确认违约概率的大小是否发生逆转的方法。

但是，我们必须注意如果发生不相吻合的情况，并非仅仅意味原因在于评级的精确度方面。按照不同的评级划分区，债务人数目极少，评级划分区内的债务人之间，存在着很强的相互关联时，也会使某一个评级划分区的违约概率不稳定，其结果是导致评级划分区的不相吻合状况的发生。

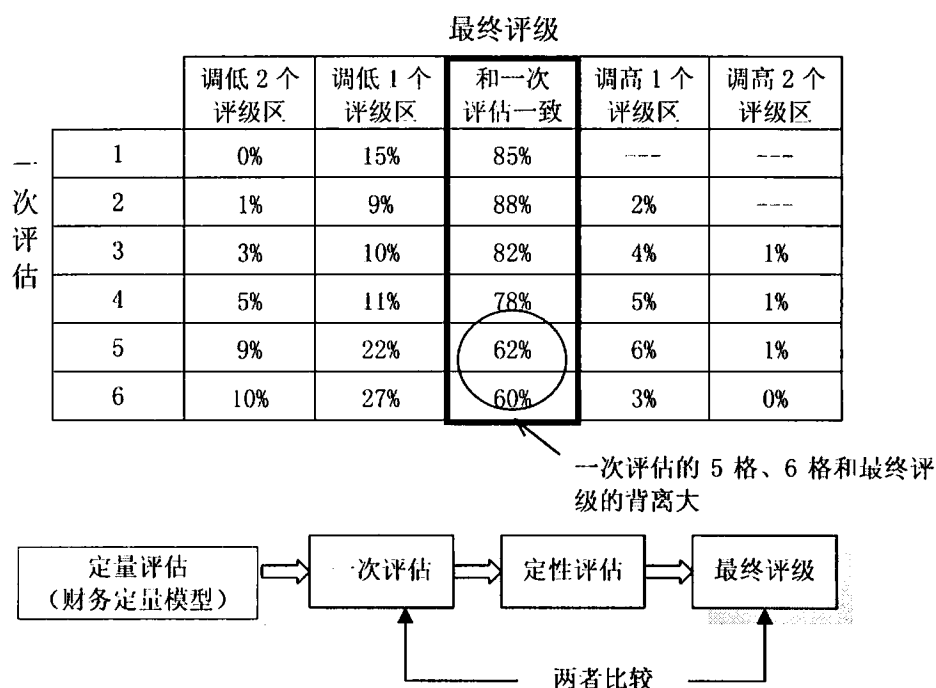
（图表 13）使用评级变化行列、不同评级违约概率进行的事后验证事例

年初评级	年末评级								违约
	1.	2.	3.	4.	5.	6.	7.	8.	
1.	83.1	12.8	2.1	0.3	0.4	0.3	0.5	0.5	0.0
2.	4.7	75.4	15.1	3.3	0.7	0.3	0.2	0.2	0.1
3.	0.2	11.9	66.5	13.9	4.2	1.5	1.0	0.8	0.0
4.	0.0	1.4	13.3	63.1	13.1	4.4	2.5	1.9	0.3
5.	0.0	0.4	4.4	24.5	44.0	15.7	6.0	4.5	0.5
6.	0.0	0.1	1.5	7.5	20.4	43.9	6.0	9.5	1.1
7.	0.0	0.0	0.5	2.8	6.8	18.9	47.8	20.0	3.2
8.	0.0	0.0	0.4	1.6	2.1	2.6	3.8	74.7	14.8

#### ④ 定量、定性评估的事后成绩

对于定量评估和定性评估相结合的评级模型，到最终评级确定之前，需要分析评级负责部门实际上做了什么样的评估非常重要。比如，最终评级结果如果同按照财务定量模型进行的第一次评估，出现大幅背离的倾向，结果导致定性调整的比重升高，这时就需要研究一下财务定量模型自身是否稳妥的问题（图表 14）。尤其是对于那些将财务定量模型作为内部评级制度核心工具的金融机构，该模型的验证及负责部门的运行评估就显得更加重要了。

（图表 14）一次评估（定量评估）和最终评级的比较验证事例



在定性评估实施得出最终评级时，也需要事后确认其是否稳妥（图表 15）。比如，只进行定量评估的一次评估结果是同一评级，但是在经过定性调整后被付与了不同评级的债务人集团，需要确认他们实际的违约率是否存在差异。相反，一次评级结果相异，但是通过分析最终评级结果相同的债务人集团的违约率，可以确认定性评估的稳妥性。

(图表 15) 定性调整后确认最终评级适当性的事例

推移前评级			推移后	
一次评估 (只是定量)	最终评级 (定性调整后)	债务人 数量	违约件数	违约率
6	5 以上	156	0	0.0 %
	6 (无变更)	223	3	1.3 %
	7 以下	147	8	5.4 %
	合计	526	11	2.1 %

} 确认违约率  
的顺序

推移前评级			推移后	
一次评估 (只是定量)	最终评级 (定性调整后)	债务人 数量	违约件数	违约率
5 以上	6	112	2	1.8 %
6		223	3	1.3 %
7 以下		117	7	6.0 %
	合计	452	12	2.7 %

} 确认有无同平均值  
大幅背离的违约率

对于一次评估【7 以下】  
有排列次序过度靠前的可能

**【专题 4】利用外部数据库及外部模型时的注意要点**

为了完善内部评级制度，有时需要利用外部数据库。那时具体的方法多种多样。比如，独自构筑评级模型时，需要企业的财务数据、违约数据等，如果自己积累的数据不够，则可以考虑使用数据库等外部数据进行补充。但是那时需要确认利用的数据是否和自己银行的授信资产组合要求相吻合。具体讲包括规模、行业、地域性等对企业的违约概率带来较大影响的内容，这就要求在自己的授信资产组合与外部数据之间不要背离较大上下工夫。

当然并不仅仅是数据，评级模型本身也可以考虑借用外部力量。比如，对于违约事例很少的优良大型企业，可以考虑利用长期按照时间顺序积累数据进行评级的评级公司建立的模型。当然在这种场合，作为前提的关于评级的思路方式（是 PIT 还是 TTC 等），模型的具体内容认真理解的同时，不要仅仅依赖模型开发者的报告，而是需要自身认真去验证和研究。这一点体现在评级模型委托外部专业公司开发时，思路也完全相同。

### (3) 评级模型的修正

参照以上评级模型的验证结果，需要判断模型是否需要修正，如果需要修正，则模型的哪一部分（财务定量模型、定性评估逻辑等）需要多大程度的修正。但是，频繁的模型修正伴随着相应的成本支出。还需要考虑确保评级模型的连贯性问题。所以，是否必须对评级模型进行修正需要认真探讨。在评级模型的运行效果已经明显退化时，需要对模型进行彻底的修正。并没有出现这种情况时，对模型进行轻微修正就可以了。

此外，评级模型的逻辑、参数（使用的财务数据）进行大幅修正，以及委托外部专业公司引入新的模型时，需要对新旧模型的评级判定结果进行比较。

## 6. 风险要素推算

在验证评级模型时，是否准确反映了债务人的违约概率是其重要的判断要点。同时，对债务人的授信信用风险评估时，除了违约概率以外，违约时的损失率和授信额度也很重要。这些评估个别授信信用风险的必要项目称为风险要素。

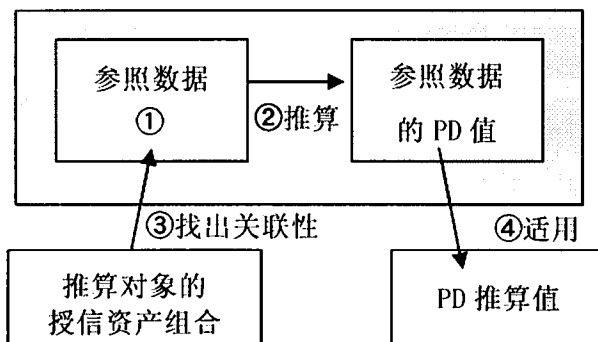
以下就 3 个基本风险要素，即债务人违约概率（PD: Probability of Default）、违约时损失率（LGD: Loss Given Default）、违约时授信（EAD: Exposure at Default）的推算方法进行概述<sup>14</sup>。

### （1）风险要素推算的基本程序

违约概率（PD）及违约时的损失比率（LGD）等风险要素的推算，基本按照以下程序进行（图表 16）。

- ① 收集参照数据
- ② 利用参照数据推算风险要素
- ③ 找出推算对象的授信资产组合与参照数据的关联
- ④ ②的推算值的适用情况及调整

（图表 16）风险要素的推算程序（PD 的情况）



<sup>14</sup> 3 个风险要素是带有不确定性的概率变数，这里的推算是算出风险要素的『期待值』的意思。



在上面内容中，③和④是在作为推算对象的授信资产组合与参照授信资产组合不同的情况时的必要程序。尤其是推算 PD 值的程序，使用内部数据和使用外部数据分开来看的话，结果如下。

(例1) 使用内部数据推算 PD 值

- ① 作为参考数据准备过去的不同评级数据（如 5 年期间的数据）
- ② 将过去按照不同评级排列数据的每年违约率单纯平均值作为不同评级的 PD 推算值。
- ③ 现在的内部评级视做同过去没有变化。
- ④ ②计算出的推算值作为现在的不同评级的 PD 值采用。

(例2) 使用外部数据推算的 PD 值

- ① 作为参考数据使用评级公司的数据。
- ② 评级公司发布的不同评级结果的违约率的长期平均值作为 PD 推算值。
- ③ 将作为内部评级参照的外部评级数据进行变换（mapping）。
- ④ 考虑违约定义等的差异，修正②的 PD 推算值。修正后的数值作为现在不同评级结果的 PD 值来采用。

包括 PD 以外的风险要素，如果详细分析推算程序，则结果如下。

① 不同评级结果的违约概率（PD）

违约概率（PD）是指，债务人在将来的一定期间内，有可能发生违约概率的计算数值。在进行该推算时，首先必须明确规定，出现什么样的情况视做违约（违约的定义）。通常，债务人的评级降到某一评级区一下时，被认为发生了违约行为。

违约的定义包括『法律上的破产』、自身审查出现的『评级实质破产以下』、『评级可疑以下』、『评级特别关注以下』等各种情况，具体采用哪个作为基准，需要金融机构根据债务人管理及风险管理理念来确定。

比如作为违约定义之一，可以如下考虑，『债权的回收已经不取决于债务人的信用状况，受抵押价值等债务人信用以外的要素影响』。现在根据这样的思路，多数银行通过自身审查将分类为『特别关注』以下作为违约对象。

除此以外，还可以有『金融机构做出该债权必须从自身资产负债表中去除判断的情况』、或者『该债权按照市场价值再评估时，做出债务人必须彻底重建判断的情况』的定义。该定义是金融机构按照自己的审查系统从分类看，属于『特别关注』以下从而被认定为违约对象<sup>15</sup>。

不同评级 PD 值的具体推算方法，包括内部数据的长期平均值推算法，同外部评级等的比较对照，以及通过统计模型直接推算等多种方法。不管采用哪种方法，样本数过少会导致统计数量的精确度欠缺，所以，参照数据的积累和完善非常重要。假如仅仅依靠内部数据不充分时，可以考虑使用外部的数据资源（共同数据库等）进行补充，那时，确认就企业所在的行业、规模、地域等属性，同内部的授信资产组合做参照对象的外部数据库的类似性分析非常重要。此外，还要注意数据库的违约定义的不同。如果定义差异较大，则需要在保守地修正外部推算值上下工夫。

## ② 违约损失率（LGD）

违约损失率（LGD）是相对于授信风险的损失预测额的比率，也可以看作『1-回收率』。作为损失额不仅要测算经济上的损失，考虑抵押等回收额之外，还需要加入回收过程中发生的费用（servicer 的服务费等工作成本）。

LGD 的推算中，也是收集关于违约债权的过去的长期数据很重要。尤其是由于景气的影响，抵押回收额等会发生变动，所以，一个景气循环以上按照长期时间序列进行的推算是最为理想的。另外，在构筑用于推算方面的数据库方面，保全的有无、抵押的种类、抵押的覆盖率、债务人信息（行业、地域、信用程度等）<sup>16</sup> 等对 LGD 带来影响的要素的收集非常重要。关于 LGD 的推算事例，我们在附件 3 进行了整理。

---

<sup>15</sup> 巴塞尔 II 中规定的违约定义是，①由于企业重组等带来的无法履行债务、②拖欠 90 天以上等被列举了出来，这个定义同我国的『特别关注』定义接近。

<sup>16</sup> 关于行业等债务人信息，基本是同 PD 相关的内容，也可能对 LGD 带来影响。比如，对属于流动性较低抵押行业的债务人的授信，出现违约时只能回收较少的额度，出现这种情况，结果是会对 LGD 带来影响。

在我国房地产抵押对债权保全所发挥的作用依旧很大。所以，在推算 LGD 的时候，准确把握房地产价格的走势、从债务人违约后到房地产抵押处分所需要的时间等很重要。迄今为止，金融机构在最终处理不良债权时，所谓的『二次损失』发生的情况并不少。这主要是对抵押房地产的价格推移、到最终处分的时间预测，过于乐观高估导致的结果。此外，房地产等的抵押价值，并非独立于债务人的信用度以外，随着债务人的破产倒闭，抵押房地产的价值也会下跌。将类似情况考虑进 LGD 的推算是很重要的。

### ③ 违约时授信 (EAD)

违约时授信 (EAD) 是违约时的授信额。贷款、债券等表内资产，其本金及帐面价值作为 EAD 是可以掌控的。但是，承诺额度等表外资产，由于将来的授信额会发生变动，所以有必要事先推算。

类似承诺限额这样的表外交易，在债务人信用度恶化的过程中，其风险有急剧增加的倾向。在我国这方面交易的数据积累还不充分，所以今后继续努力积累数据的同时，在充分的数据能够被确保之前，应该保守地进行评估。

## (2) 风险要素的查证

由于推算风险要素时需要使用过去的的数据，所以其推算值是否稳妥准确需要持续进行验证。验证的方法各种各样，比如对于 PD 可以考虑使用以下的方法。

### ① 推算值的顺序

不同评级的 PD 值按照评级级别是否整合在一起，即需要最低限度地确认评级级别间推算的 PD 大小是否发生了逆转。

### ② 逆向测试

一般的验证方式是将推算的 PD 值同实际的违约率进行比较，也就是所说的逆向测试。然而，同每天都可以取得数据的市场风险不同，通常 PD 是以年为基础计算的，所以，有时无法确保足够长时间按时间序列的数据，从统计角度看无法进行有效的验证。

### ③ 按时间序列看稳定性

如前论述，从 PIT 的立场看，不同评级的实际违约率无关于景气局面的变化，稳定地进行推移。所以，采用 PIT 法时，通过验证这一点，可以评估 PD 推算值的妥当性。而另一方面，从 TTC 的立场出发，即使是相同的评级，每年的实际违约率随着景气循环而变动。所以，采用 TTC 法时，PD 的推算值需要预留一定的变动区间，从而验证实际违约率是否收在预留的区间之内。假设如果没有收在预留的区间内，则①由于景气预想外的变动导致的、②样本数值不够出现的异常数值、③评级、PD 推算本身缺少妥当性，必须认真分析其具体原因，如果是③的情况，则需要修正评级模型。

#### 【专题 5】关于逆向测试的统计验证手法

逆向测试时的统计验证手法包括二项审定 (binomial test)、正规审定 (normal test)。二项审定是针对某评级区的债务人，实际观测到的违约率，验证其与事先推算的 PD 值之间的误差是否在规定的范围内。比如，评级区的债务人数目是 1000 个，假设事先推算的 PD 为 1%，则平均有 10 个 ( $1000 \times 0.01$ ) 违约发生，而实际情况是有可能超过 10 个或低于 10 个。评估类似上述实际的违约项目数是否在一定的概率上有意地发生就是二项审定。

而正规审定是假设不同的评级 PD 按时间序列是一定的，审定过去实际观测到的违约率值是否适合上述一定的数值。比如假设推算的 PD 为 1%，过去 5 年的实际违约率为 1.2%、1.8%、0.8%、0.7%、1.1%，出现同 1% 背离的结果时，评估该结果是否为有意地引起的。

但是，在没有足够的样品时，这种审定的可信程度就会降低。而且，违约经常是债务人之间有较强相关联的情况下发生的，而上述审定进行的时候，多数是以违约独立发生作为前提的，所以，必须注意并非所有的审定都是很可靠、确切的。

**【专题 6】如何事后确定评级实施的特征（是 PIT 还是 TTC）？**

如上所述，PIT 和 TTC 的区别在于，景气变动给债务人带来的影响是通过评级变化予以吸收（PIT），或者被不同评级的实际违约率变化予以吸收（TTC）。所以，将应对景气循环的授信资产组合的质量变化，通过分成评级的变化说明部分（PIT 模式）和不同的评级实际违约率变化说明部分（TTC 模式），对评级实施的特征进行事后评估。

最近，通过日本银行的考查，对内部评级制度比较完备金融机构的不同评级实际违约率的变化进行验证的结果，表明同景气循环是相关的。这也说明这些金融机构的内部评级制度，虽然有程度上的差别，但可以解释为拥有 TTC 的特征。

当然，PIT 部分和 TTC 部分划分开来看，既有介于两者的中间型，也有和 PIT 相近的金融机构。此外，在有的金融机构，个别评级对景气循环的反映也不同。也就是说，高位评级的违约率对景气的感应度也相对较高（TTC 模式），低位评级的感应度则相对较低（PIT 模式）。正因为有这样的事例，在实际验证 PD 推算值的准确性、稳妥性时，需要考虑到上述要点。

## 7. 内部评级制度的灵活运用

内部评级制度是强化信用风险管理，合理、高效的授信运营，以及制定经营战略的基础。但是，如果引入的内部评级制度仅仅止步于形式内容，在实际的风险管理及经营当中没有被充分灵活运用，则制度本身的引入就会失去意义。所以，在完善内部评级制度方面，需要探讨其灵活运用的方式方法。

内部评级制度的具体的灵活运用方法，首先可以作为营业、审查部门个别授信的风险认识，以及定价等的工具来使用。之后中间部门运用内部评级、评级变化行列、不同评级的 PD 等内部评级制度的产出信息<sup>17</sup>，开展授信资产组合的监管跟踪（把握不同评级的余额），设定各个评级的定价实施规则，信用风险计量，信用成本估算等工作。

作为参考，日本银行最近的考察中观察到的内部评级信息，各个评级 PD 运用事例，如下所述。

### ① 内部评级信息的运用事例

（授信执行阶段）

- 设定各个评级的授信限额。比如，评级低的企业限额设得较低，事先减小授信集中风险。
- 设定各个评级的授信决定权限。比如，评级越低营业店的决定权限授信限额越低。
- 简化优良评级企业的审查程序。有助于据此削减整体的审查成本。

（中间管理）

- 针对不同的评级对授信企业的管理。比如，对于『正常』级企业不是采取完全统一的对策，而是对评级发生降低的企业及相对处于低位的企业，强化对其的跟踪监管。或者尽早确定低位评级企业的应对方针，在问题明确显现出来之前找出对策。
- 把握授信资产组合整体的信用状况。比如，通过跟踪调查评级的变化行列、行业及各地区的余额变化状况，发现品质低劣、比较醒目的评级区、行业及地域等。

---

<sup>17</sup> 内部评级制度的产出信息，要求最大程度地在金融机构内被综合统一运用。此外，如果发生任何修正的情况，需要事先对其稳妥性进行验证。比如，利用各个评级 PD 的信息时，定价上的 PD 和信用风险测量上的 PD 出现差异时，其理由必须予以明确。

## ② 各个评级 PD 的运用事例

- 信用风险的计量和资本配置。将各个评级的 PD 作为信用风险计量模型的输入信息利用（关于信用风险计量请参照第 9 章）。并根据测定的风险大小向各部门摊入资本。
- 考虑信用风险定价（设定贷款基准利率）。贷款基准利率是根据每个授信项目不同，在筹作利率和经费比率的基础上加上信用风险差幅（信用成本率与资本成本率之和），再加上目标收益比率最终确定。在这当中，信用成本率是依据各个评级的 PD 等进行推算的。此外，以各个评级 PD 为基础信用风险大小被推算出来，和此相对应的资本摊入执行的时候，表明在资本成本率的计算当中也使用了 PD。如此，考虑信用风险因素在内的定价确定后，可以有效防止潜在的不合理、不合算贷款，在没有被明确认识的情况中任意扩大。
- 把握贷款的经济价值<sup>18</sup>。使用 DCF（Discounted Cash Flow）评估贷款债权的经济价值时，从刨除信用风险的未来的现金流当中，计算现在的贴现价值，该信用成本的推算需要利用各个评级的 PD 等。

---

<sup>18</sup> 请参照日本银行 2003 年 4 月发表的「把握贷款的经济价值和其意义-面向金融机构、企业的经营模式变革」。

## 8. 内部评级制度的验证

金融机构所处的环境在不断地发生变化，在此当中，为了维持内部评级制度的有效性，需要持续不断地验证其确切性和稳妥性。我们已经在第5章、第6章对评级模型和风险要素推算的验证方面，以技术侧面为中心进行了整理。在这里，我们再从更广的视野，对内部评级制度的整体验证思路进行分析（图表17、图表18）。

内部评级制度的验证当中，以下几点尤其重要。

### ① 经营者对验证的参与

经营者必须理解内部评级制度和其验证的必要性，并且需要积极参与验证体制及验证工作框架的制定。

### ② 信用风险综合部门（中间部门）和授信监察部门双方的验证

授信规划部等信用风险综合部门（中间部门），负责内部评级制度整体的规划方案、规则手续的制定，同时开展验证工作。除此之外，授信监察部门等从第三者的立场，另行开展验证工作，并且需要将结果汇报经营者。

### ③ 明确验证的框架原则

明确验证工作的框架原则和指导方针的基础上，将其内容传达给组织内所有人员。

### ④ 持续验证的实施

内部评级制度的验证，并非仅仅体现在引入的时候，引入之后也需要继续执行，如果有必要还需要进一步改善。验证的频度根据不同的对象虽然有区别，但是对评级制度整体带来较大影响的重要事项，需要以较高的频率进行验证。

### ⑤ 制度和运用的双方验证

验证内部评级制度体系及评级模型的准确性的同时，还需要对营业部门和审查部门在实际工作中，如何运用制度的情况进行验证。

### ⑥ 对验证方式持续不断的研究和修正

验证手法并非一成不变的，需要持续不断地研究和修正。尤其是使用统计手法的验证，因为需要依靠大量可以利用的数据，所以在数据大量累积起来的时候，需要在采取新的验证手法等方面下工夫。

### ⑦ 验证时所必须的数据的准备



为了实施内部评级制度验证，需要准备评级模型中债务人的财务数据和定量评估、定性评估中反映评级变化的数据（参照专题 7）。随着数据的累积，可以确认财务定量模型和定性评估的成绩。此外在评级模型发生变更的时候，需要追溯到过去对评级重新执行。

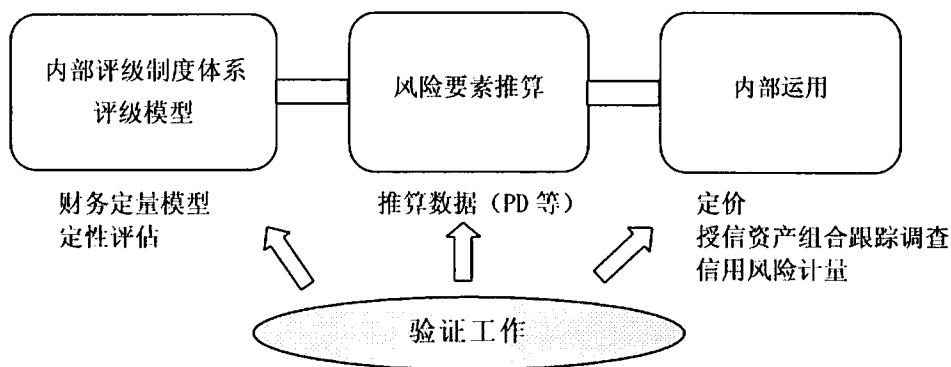
⑧ 验证有关信息的文字化

有关验证手法和结果的相关信息需要以文字方式保留。

（图表 17）验证要点

验证的目的	确保内部评级制度的确切性 提高信用风险评估的准确度
验证对象	内部评级制度体系 债务人评级和项目评级（或者 LGD 评级） 评级模型（财务定量模型等） 数据准备状况 运用状况、评级程序 （营业网点、审查部门是否在正确实施评级） 各种产出（评级变化行列、PD、LGD 等） 产出的使用状况（授信资产组合跟踪调查、定价、对经营者的汇报）
验证体制	信用风险综合部门 授信监察部门 经营者
验证手法	现场验证和非现场验证 定量验证（统计手法）和定性验证 各种文件的检查 模型检查（计算机程序） 样品检查（确认运用状况） 标准化检查（同外部资源进行比较等） 逆向测试（推算值和实际值的比较）
验证时间	定期实施 （但是频度由验证目的及实施对象的重要程度决定）

(图表 18) 验证工作对象

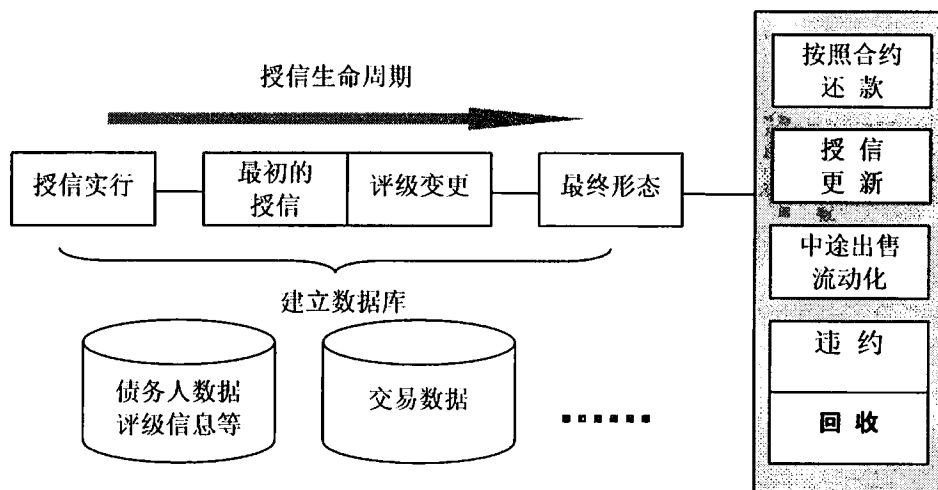


### 【专题 7】内部评级制度的数据准备

为了提高内部评级模型及风险要素的精确度，累积评级时所需要的数据是非常重要的。尤其是为了使第三者能够准确、高效地开展验证工作，必须准备详细且可信度高的数据。

比如，通过累积债务人过去的评级履历、定量及定性评估的结果，则可以追溯过去验证各个评级的 PD 预测能力。此外，开展评级时有关财务及债务人特性的数据，在内部评级制度发生本质变更（特别是评级模型变更），金融机构合并重组构筑新的制度时，上述数据会发挥重要作用。

收集债务人及有关交易数据时，通过自授信开始到授信终结期间（授信的「生命周期」）的观察，应该将最初授信及其变更履历、最终处理的相关信息长期累积。尤其在构筑数据库时，事后进行一些多样、高效的分析是非常有必要的。



需要收集的数据项目如下：

（关于债务人的信息）

债务人评级、评级实施/更新日期、至最终评级的评估（一次评估、二次评估等）  
定量信息（财务数据）、定性信息（行业、企业特性等）、外部信息（外部评级等）

（有关交易信息）

项目评级、授信额、现金流信息、承诺限额  
授信日的、授信形态、优先/劣后状况、抵押种类（货币资金、国债、房地产等）  
抵押价格和评估日期、抵押覆盖比率，有无担保

（有关违约信息）

最终处理形态、违约日期、违约状况（法律处理、拖欠、条件缓和等）  
违约时风险  
回收额和回收日期、回收手段（抵押、担保等）、回收费用  
将回收额贴现到违约发生时的贴现率

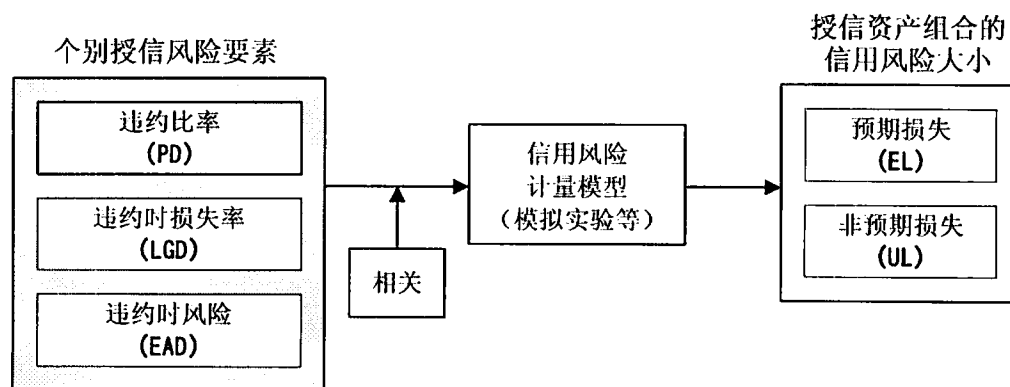
## 9. 信用风险的计量化

到这里本报告已经介绍了内部评级制度核心内容的金融机构信用管理问题的各种论点，随着内部评级制度的逐步完善，可以通过定量分析把握金融机构的授信资产组合持有的信用风险。最后对于信用风险的计量问题，如下将我们的思路整理一下。

### (1) 预期损失 (EL) 和非预期损失 (UL)

如果已经预测到的风险要素可以推算，则可以推测出授信资产组合整体包含的信用风险的大小<sup>19</sup> (图表 19)。通常信用风险依据预期损失 (EL: Expected Loss) 和非预期损失 (UL: Unexpected Loss) 计量。预期损失 (EL) 是在一定的授信期内 (比如 1 年)<sup>20</sup> 预期发生的损失额的平均值。而非预期损失 (UL) 可以定义为，以现有的授信资产组合结构为前提，按照一定的概率推算产生的最大损失<sup>21</sup> 中，扣除 EL 之后的额度。所说的一定概率是指假设的 99%、99.9%，该概率有多高被称为信赖水平 (图表 20)。

(图表 19) 授信资产组合的信用风险计量 (设想)

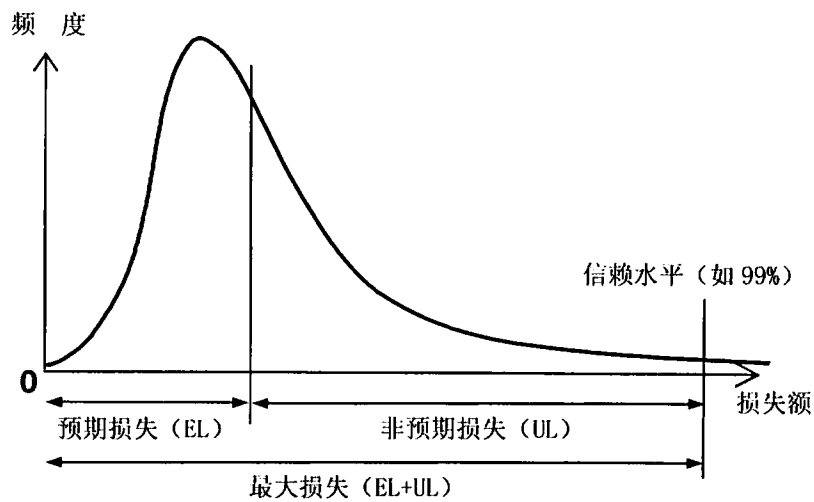


<sup>19</sup> 信用风险的测算，包括①债务人只有在违约时视做发生损失的破产方式 (default mode)、②除违约以外，再加上债务人信用度 (评级) 降低带来的经济价值减少，考虑这种损失的市场价值方式 (mark-to-market mode) 这两种。

<sup>20</sup> 持有期限 (风险评估期) 一般确定为同业务运行周期相吻合的 1 年，或者定为授信合约上的剩余期限。即使合约体现的是短期贷款，由于采取滚动式信贷实质上相当于长期信贷的情况时，剩余期限如何设定是一个重要的论点。

<sup>21</sup> 「一定概率 (如 99%) 推算产生的最大损失」的意思是「在此之上损失发生的概率为 1% (100%-99%) 的损失额」。

(图表 20) 授信资产组合的损失分布 (设想)



而且需要补充说明的是，授信期以 1 年为前提的信赖水平 99% 的最大损失 (EL+UL) 意味着 100 年期间只发生一次超过该损失的情况。

通常，EL 是为了抑制风险，计算平均发生所需的成本，用呆帐准备金拨备，而 UL 是用自有资本应对该风险带来的潜在损失。

另外持有的授信资产组合，如果贷款集中投向一部分企业，出现由于行业集中、地域集中、连环破产可能性较大的情况时，贷款对象之间的违约关联程度也会加强。出现这种情况时，同 EL 对比看到的 UL 会出现变大的倾向（图表 20 损失分布的右侧坡度变厚）。相反，债务人数非常多呈现小额分散化趋势时，或者违约相关程度降低时，同 EL 对比看到的 UL 会出现变小的趋势<sup>22</sup>。

<sup>22</sup> 这可以用简单的数值举例说明。假设现在的整体授信额度是 1000 亿日元，个别授信的年 PD 一律为 2%、LGD 一律为 100% 的假想授信资产组合。

(事例①) 授信组合仅由单一授信 (债务人 A) 构成的情况。因为 A 的 PD 是 2%，今后 100 年间按照概率看，2 年违约，其损失额是 1000 亿日元的全部损失 (剩余的 98 年，由于 A 不会违约所以损失额是零)。所以，可信水平 99% 的最大损失额 (EL+UL)，及 100 年当中第 2 大的损失额为 1000 亿日元。

(事例②) 授信资产组合由 2 个授信 (债务人 A 和 B) 构成，授信额度共计 500 亿日元。A 和 B 的违约关联为零 (不相关)，今后 100 年当中，第 2 大的损失额是 A 或 B 破产时的 500 亿日元 (两者同时破产产生 1000 亿日元的概率为， $2\% \times 2\% = 0.04\%$ ，不满 1%)。

如上所述，授信资产组合的小额分散构成越多，信赖水平 99% 的最大损失额就会变小。但是，事例② A、B 之间的相关如果为 1 (完全相关)，则同事例①相同，最大损失额达到 1000 亿日元。为什么会这样呢？因为 A 破产时，B 也会破产。也就是说，违约相关程度越强，最大损失额度越大。

## (2) 风险计量

每个债务人的 EL 可以根据第 6 章论述的 3 个风险要素的预期值相乘算出来<sup>23</sup>。而且，授信资产组合整体的 EL，是各债务人的合计。

$$\text{预期损失 (EL)} = \text{EAD} \times \text{PD} \times \text{LGD}$$

UL 的计算在上述 EL 计算时使用的 3 个风险要素预期值之外，还需要加进这些概率的分布形状、债务人之间的违约相关因素等。所以，从技术层面讲，比 EL 的计算更为困难，其计算方法多采用蒙特卡洛模拟实验。

### 【专题 8】巴塞尔 II 的风险比重函数

巴塞尔 II 的第 1 核心（所需要的最低自有资本）中的内部评级手法，是通过风险要素推算信用风险大小的，并且要求使用巴塞尔 II 的计算公式（风险比重函数）。风险比重函数是在授信集中程度、违约相关性等一定的前提下，如果 PD、LGD、EAD 等的风险要素齐备，就可以计算风险大小。

另一方面，正如要求作为第 2 支柱的另行评估授信集中风险一样，如果仅仅依靠第 1 支柱的风险比重函数，金融机构面对的风险有可能会有遗漏的地方。所以，金融机构应该在准确认识自身风险的基础上，推算出和其相对应的信用风险的大小，尤其是 UL 值。

同时，作为第 1 支柱的风险比重函数，理论结构是以一个要素 Merton 模型为基准。该模型关于债务人企业价值的变动，是通过一个宏观层面的共性要素（systematic factor）和企业固有要素（idiosyncratic factor）的加重和来表现的。这两个要素时时刻刻在发生变化，企业价值也随之变化。当企业价值降低到一个被认可的数值时，则可以视做会发生违约。比如将景气变动看作 1 个宏观的共性要素，则个别债务人的违约行为，同个别债务人的特性要素一同，通过景气的变动进行说明。受景气变动的影响越强，更多的债务人会同时受到或大或小的影响，个别债务人和授信资产组合整体的 UL 就会变大。计算 EL、UL 的金融机构，大多利用同样的要素模型。

<sup>23</sup> 这里假设 3 个风险要素是相互独立的概率变数。

### (3) 压力测试

根据信用风险计量模型算出的风险大小，是在一定的前提条件下推算出来的以外，如果无法确保模型及风险要素数据的稳定性，则该推算值会时刻伴随不确定性。所以，为了开展更有实效的风险管理，需要准备应对上述不确定性的手段。压力测试是为了应对有一定发生可能性的市场环境发生急剧变化时的危机事例（压力事例），借此分析其带来的影响。实施测试的时候，需要判断压力事例发生的情况下，授信资产组合的运营方针是否稳固，对经营能力是否充分进行评估。作为压力事例可以考虑到的有经济增长率的下滑，股价、房地产价格的大幅下跌，大额贷款企业的经营状况恶化，风险要素的大幅变动等。设定该事例时，认真参照自身的授信特性（比如债务人所处的行业、规模等）非常重要。

在决定授信资产组合的运营方针时，希望压力测试的结果能够被经营者及风险管理部门、营业部门作为研究材料积极运用。为了促进经营者充分意识到压力事例来设定自有资本缓冲，找出应对的策略，要求中间部门提供有说服力的压力事例。

#### 【专题 9】对特殊表内资产的压力测试

作为压力测试的一个事例，让我们观察一下授信额（EAD）将来有可能增加的特殊表内资产。

如上所述，类似普通的贷款资产等表内资产，其贷款本金在贷款期内基本并认定为不会变化，所以本金金额、帐面价值等多被直接当作 EAD 考虑。但是，即使在个别贷款的贷款约定上本金是确定的，由于金融机构和债务人之间的特殊关系，发生违约之前，也会出现债务人授信额度扩大的情况。比如，金融机构 A 是大额债务人 B 的主银行，在 B 的经营恶化时，A①违约时对自身经营带来的影响大小、②作为主银行的道义责任（为规避声誉风险的显现）等角度出发，有时会替代其它银行提供信贷。就是所谓的『风险向主银行集中』的现象。实际上，许多金融机构对经营恶化的大额客户的贷款，该客户从违约或放弃债权开始，一直到再建结束，贷款增加的情况更多。

当然，这类的『风险向主银行集中』产生的债务人特性、随着『风险向主银行集中』而增加的贷款估算也会很困难。但是，从迄今为止我国的金融经验看，『风险向主银行集中』的风险（换言之就是作为主银行风险）某一时期会极度增

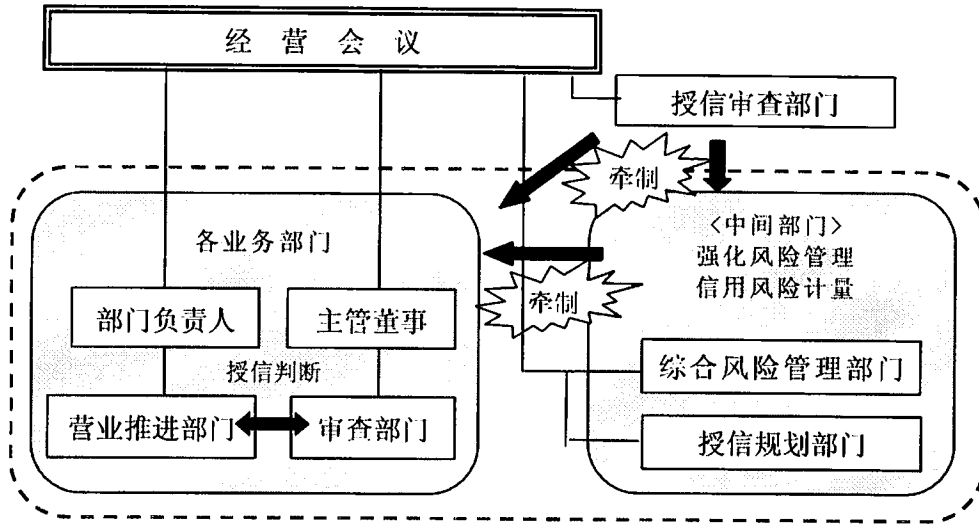
大。所以，在使用压力测试时，要尽可能地将这种风险考虑进去。

以 上



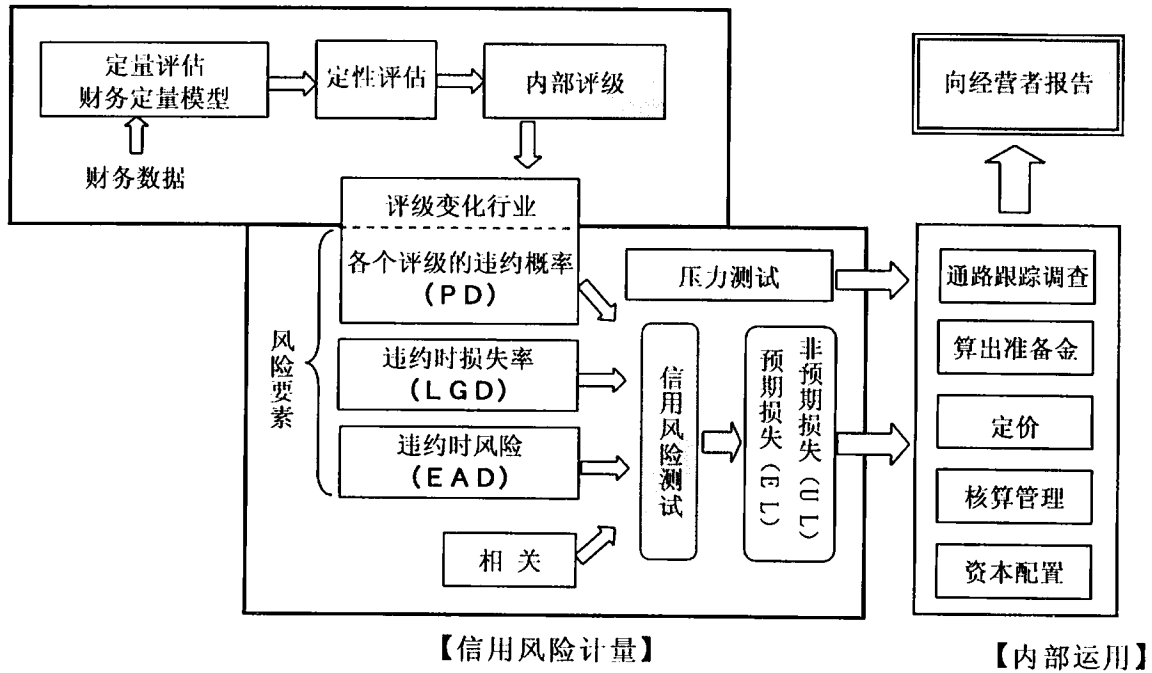
(附件 1) 信用风险管理体制和强化信用风险管理设想

信用风险管理体制设想



强化信用风险管理设想

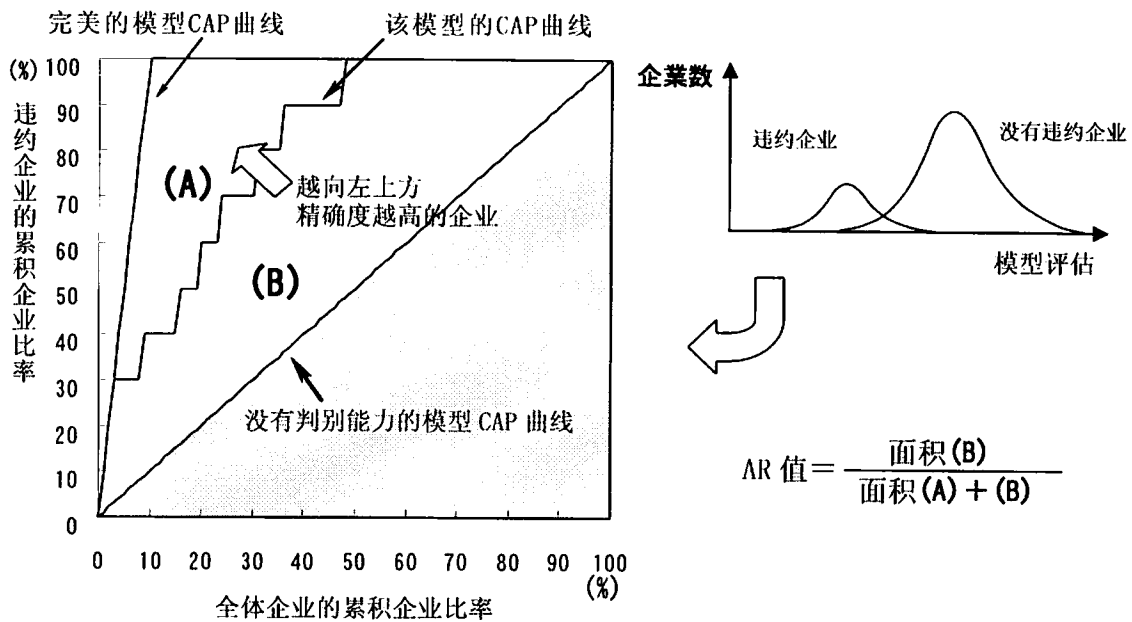
【内部评级制度】



(附件 2) 为实施验证的统计数量事例

(1) CAP (Cumulative Accuracy Profiles) 曲线和 AR (Accuracy Ratio) 值

作为评估评级模型精确度的方法,可以画出CAP曲线(也称为Gini Curve、Power Curve、Lorenz Curve),从而通过定量统计AR值(也称做Gini系数)计算模型运行效果的方法。这种方法通过排列模型评估的债务人信用度得分,比较实际违约债务人的数量,评估模型的判别能力。

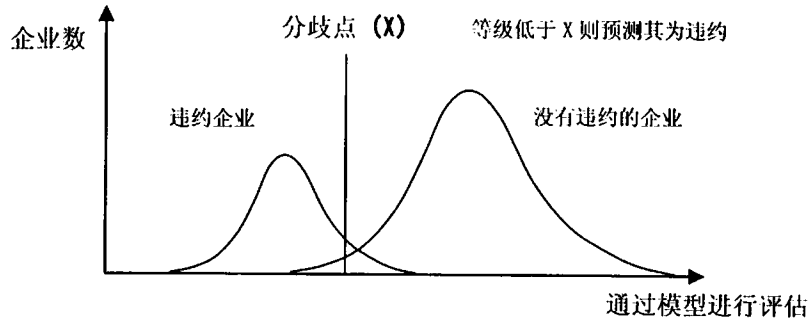


用于推算的债务人数为  $N$ , 其中实际违约企业数假定为  $n$ 。CAP 曲线是横轴模型中信用度得分低的前  $x$  企业占整体的比率 ( $x/N$ ), 纵轴表示模型信用度得分低的前  $x$  企业当中, 实际违约  $n_x$  企业占违约企业的比率 ( $n_x/n$ ), 来描述两者关系的。

CAP 曲线, 判别能力高的模型位于左上方, 完全没有判别能力的随机评估显示模型位于 45 度直线处。45 度线和 CAP 曲线之间区域面积算出的就是 AR 值。其面积越靠近完美的评级模型区域, 即 AR 值越高, 评级模型的精确度就越高。

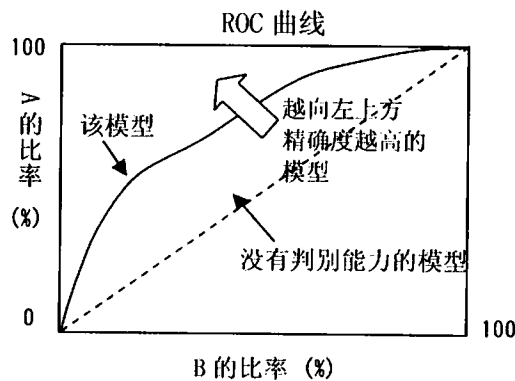
(2) ROC (Receiver Operating Characteristic) 曲线和 AUC (Area under the Curve) 值

评级模型的评估显示，低于任意分歧点 X 是违约，高于分歧点的则是非违约部分。通过这种方式定义的模型做出的违约和非违约预测，和实际的违约企业及非违约企业进行比较，也可以对模型的判别能力做出评估。



如下表所述，预测值同实际数值对照会出现 4 种情况。模型的判别能力越高，表中的 A 和 D 区域的比率就一定越高。ROC 曲线在分歧点 X 的位置，从得分差的地方向好的地方移动（上方的图标从左向右），预测准确的区域 A 的比率= $A / (A+C)$ ，预测错误的区域 B 的比率= $B / (B+D)$ 。判别能力高的模型其 ROC 曲线位于左上方，完全没有判别能力的随机评估模型位于 45 度直线处。还有，AUC 值可以定义为 ROC 曲线以下领域的面积。

		实际成绩	
		违约	非违约
模型预测	违约	A (预测准确)	B (预测失误)
	非违约	C (预测失误)	D (预测准确)



### （附件 3）违约时损失比率（LGD）推算事例

我们考虑使用过去的内部数据（比如 7 年中违约债权的回收信息）作为参考数据的情况。数据按照『何种属性的债权在违约时会带来什么程度的经济损失』的方式来整理。为了评估经济损失，抵押、担保人的回收额、回收时间、回收所需要的费用等数据也是必须的。

$$\begin{aligned} \text{LGD} &= 1 - \text{回收率} \\ &= 1 - (\text{回收额} - \text{回收费用}) / \text{违约时授信额度} \end{aligned}$$

#### ① 收集 LGD 数据（建立参照数据库）

必须收集对 LGD 有影响的以下数据

- 有无抵押
- 抵押种类（金融资产、房地产、其它的实物抵押等）
- 抵押覆盖率
- 有无保证
- 债务人信息（行业、地域、信用度等）<sup>24</sup>

#### ② 违约债权分类

下面按照主要要素对违约债权进行分类。然后，分别算出分类的平均 LGD 值（下面的③、④进一步论述），并将该数值作为分类整体的推算值。比如，抵押的保全程度（抵押覆盖率：高~低），根据不同的抵押保全状况，平均 LGD 的区分如下表。

		抵押分类	
		房地产抵押	金融资产抵押
抵押覆盖率：高	x %以上	LGD=A%	LGD=B%
中	y %~ x %	LGD=C%	LGD=D%
低	y %以下	LGD=E%	LGD=F%

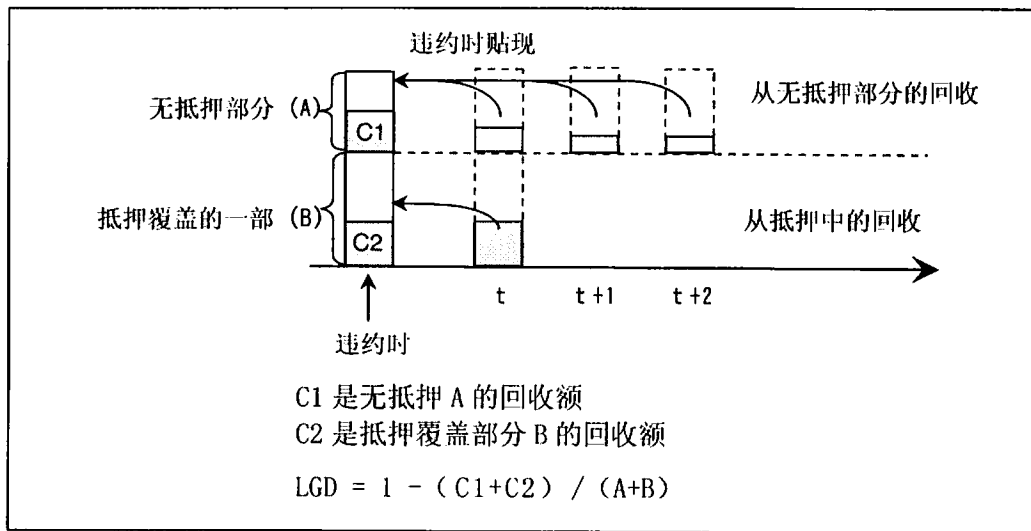
在这里将数个要素结合在一起设定好评级区分，按照不同的评级区分也可以推算出 LGD 值。依据上面的例子，以对应抵押覆盖率和抵押种类结合的 LGD 为基础，设定评级区分。当然，使用各种各样的说明变数（抵押覆盖率之外，还包括抵押种类、债务人所在行业等）可以让 LGD 回归，利用得到的方程式可以直接推算每个项目的 LGD 值。但是，违约债权的项目数是有限的，所以如果过度增加评级区分和说明变数，则有可能得不到统计上有意义的 LGD 推算值了。

<sup>24</sup> 债务人信息基本是对 PD 带来影响的信息，但是也需要考虑对 LGD 带来影响的情况。

### ③ 经济损失测算

LGD 推算违约定义，应该和 PD 推算定义相同。此外，由于 LGD 不是会计上的损失，而是经济损失，所以对于该损失的测算，抵押及担保回收额之外，还需要考虑回收费用的问题<sup>25</sup>。当然，各期的回收额，通过反映回收不确定性的贴现率，在违约时退还的基础上用来计算 LGD。

违约债权 LGD 的计算例子



### ④ 依据长期数据推算 LGD

过去所有的违约债权，通过上述③计算出 LGD，然后求其平均值。此外，由于 LGD 受景气循环的影响，所以尽可能地准备长期数据，将景气循环等影响充分考虑后，再计算 LGD 数值。

<sup>25</sup> 考虑回收额及回收费用预算时，需要注意以下要点。

- 长期回收的想法（不考虑违约后经过一段时间以后的回收）
- 回收费用的范围
- 从发生违约到回收之间的贴水情况（在此期间时间价值的测算如何选择）