鹏元國際 PENGYUAN

# 应用违约数据和量化指标评估信用评级质量

## 目录



## 联系我们

**主分析师**

姓名 陈科，PhD
职位 首席评级官
电话 +852 3615 8316
邮箱 ke.chen@pyrating.com

**分析师**

姓名 王诗雨，PhD
职位 分析师
电话 +86 755 8287 1237
邮箱 shiyu.wang@pyrating.com

## 观点提要

信用评级是以第三方的立场，根据独立、客观、公正的原则，就债务人在未来一段时间依约偿债的意愿及能力做出的判断，是对债务人的信用质量提供的前瞻性观点。信用评级作为债券市场发展中的基础设施之一，发挥着揭示信用风险、辅助市场定价、提高市场效率等重要作用。然而，信用评级只有在确保评级质量并得到市场参与各方认可的情况下才能有效发挥作用。要提高评级质量和认可度，我们需要首先确立一套评价评级质量的方法和指标，从而为市场参与各方对评级质量的判定提供合理依据。

目前，中国信用评级准确性检验采用较多的方法是信用利差检验。该方法的广泛使用主要是因为中国债券市场历史违约数据不足，无法采用依赖违约数据的检验方法直接评价信用评级是否准确度量了信用风险，且该方法在假设和检验准确性上都存在一定缺陷。为了更准确的评判信用评级的质量，市场需要一套更合理、精准、并能直接反映度量信用风险准确性的信用评级质量评价指标。特别是近年来中国债券市场违约数据的不断积累，利用违约数据来评价评级质量的量化检验方法应该得到更多的应用。

本文通过借鉴国际上的评价方法，总结并构建了评级质量的量化评价体系，希望为中国的评级机构、投资人、发起人和监管部门等各方提供更丰富有效的评级质量检验工具。我们构建的量化评价体系分为三个维度——信用排序能力、预测准确性和评级稳定性：

- **信用排序能力**代表的是区分信用质量好的主体和信用质量差的主体的能力。考察的是信用评级度量相对信用风险的准确性。评估信用排序能力的常用的指标有三个: AUROC(Area Under the Receiver Operating Characteristic Curve)、 AR(Accuracy Ratio)和 K-S 统计量(Kolmogorov-Smirnov Statistic)。
- **预测准确性**衡量的是评级对各信用等级对应的违约率预测的准确性，一般是通过观测评级等级的累计违约率和预期违约率或理想违约率之间是否一致来判断，一般使用统计学中的假设检验来完成评估。常用的检验方法有二项式检验，卡方检验和正态分布检验。
- **评级稳定性**的内容包括评级结果的稳定性和等级对应的违约率的稳定性，主要通过观测评级方法得出的实际信用等级分布和预期信用等级分布之间的差异来评估。常用的检验方法有对级别迁移率的分析和计算群体稳定性指标(PSI)。

我们使用构建的量化评估方法对中国主要七家信用评级机构的企业评级的评级质量进行了评估。

在信用排序能力评估中，我们衡量七家评级机构的评级能否有效区分一年内将违约的发行人主体。AUROC、AR 和 K-S 指标的估算结果显示在信用排序能力评估中表现较好的三家机构为中债、中证鹏元和中诚信。

在预测准确性评估中，我们衡量各家评级机构在不同信用等级下的一年期违约率是否符合预期。检验结果显示，预测准确性较好的三家机构为中债、中诚信和中证鹏元。

在评级稳定性评估中，我们通过构造评级迁移矩阵、计算调整率和调整幅度来分析评级调整情况以及衡量其评级稳定性。根据调整率等指标显示，稳定性较好的三家机构为中证鹏元、中诚信和东方金城。

## 介绍

信用评级是以第三方的立场，根据独立、客观、公正的原则就债务人在未来一段时间依约偿债的意愿及能力做出的判断，是对债务人的信用质量提供的前瞻性观点。信用评级在金融市场运行中发挥着解决信贷及债券市场参与方信息不对称、揭示信用风险、辅助市场定价和提高市场效率等作用，是金融市场重要的基础设施。

目前中国规范评级行业的监管依据是 2019 年底中国人民银行、国家发展和改革委员会、财政部、证监会四部委联合发布的《信用评级业管理暂行办法》，该办法的发布意味着中国信用评级业正式进入统一监管时代，将极大地促进中国信用评级业的规范发展。同时，中国银行间市场交易商协会和中国证券业协会每年会联合发布信用评级机构业务市场化评价结果，旨在推动信用评级管理的协调统一，加强对债券市场信用评级机构的自律管理，促进信用评级行业规范发展和评级质量的提高。发改委也会每年开展企业债券主承销商和信用评级机构信用评价工作，以此规范企业债券主承销商和信用评级机构相关业务行为，提高承销和信用评级质量，加强企业债券市场信用体系建设和事中事后监管，推进企业债券市场健康、可持续发展。

在信用评级业的监管和自律管理中，对信用评级机构的评级质量进行评估和检验是十分重要的一环。中国监管目前主要通过计算实际违约率、级别迁移率和利差分析来对评级质量进行验证[1]，还未出现系统的评级质量检验体系。此外，通过信用利差来检验信用评级质量的方法也存在着比较明显的缺陷。首先，观测到的信用利差不仅体现了债券的信用风险溢价，也同时包含了流动性风险溢价，我们在实际检验中并不能有效剔除信用风险以外的影响因素。其次，利用信用利差这类市场信息的前提假设是市场有效性——债券价格应该充分反映所有相关的信息，然而目前的债券市场并非完全有效。因此，这类对评级质量的间接检验方法无法准确直接的评估信用评级是否准确的度量了信用风险。近年来，随着中国债券市场上违约数量的增加，中国的评级行业和监管应该更多的利用违约数据来直接评价信用评级的质量，进而帮助市场各参与方更好的理解信用评级，促进评级机构不断改进评级方法。

本文借鉴国际上先进的对评级质量评估和检验的经验，基于欧洲证券和市场管理局（ESMA）发布的《信用评级机构方法有效性评估和审查指南》[2]中对评级质量的评估检验指引，总结并构建了评级质量的量化评估体系，希望能为中国的评级机构、投资人、发起人和监管部门等各方提供更多丰富有效的评级质量评价工具。我们构建的量化评估体系分为三个维度——信用排序能力、预测准确性、评级稳定性，下面我们将对这三个维度涉及的量化方法和指标进行介绍，并用其对中国评级机构的评级质量进行合理评估。

## 信用排序能力

排序能力 (Discriminatory Power) 代表的是评级在一定时间范围内按照未来状态（违约或未违约）对所评对象进行排序的能力，简单来说，是有效区分信用质量好的对象和信用质量差的对象的能力。信用排序能力是评价评级质量最核心的指标，因为信用评级本质上是对信用风险的一种相对排序。信用排序能力最常用的度量指标有以下三个：

- AUROC (Area Under the Receiver Operating Characteristic Curve)；
- AR (Accuracy Ratio)，也称为基尼系数 (Gini Coefficient)；
- K-S 统计量 (Kolmogorov-Smirnov Statistic)。

下面我们将对这三个指标进行详细介绍。

[1] 《银行间债券市场非金融企业债务融资工具信用评级业务信息披露规则》第十八条

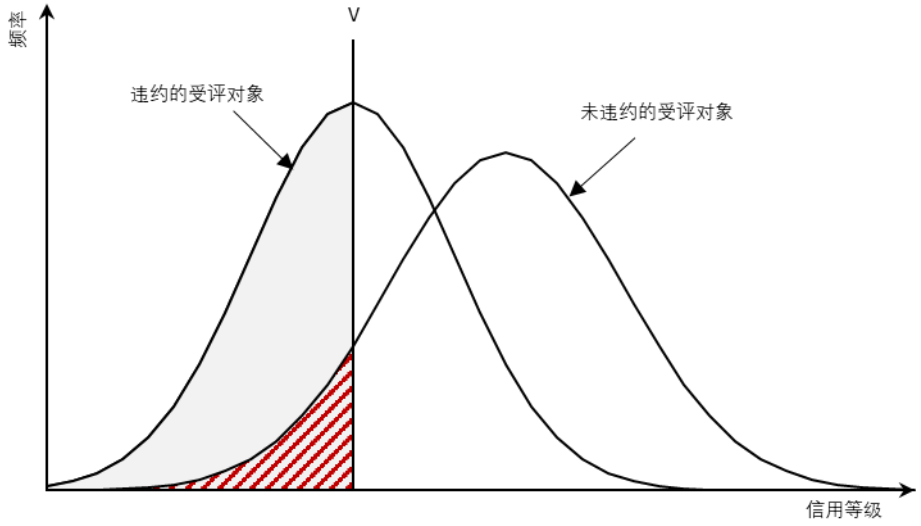
[2] “Guidelines on the validation and review of Credit Rating Agencies’ methodologies”ESMA, 23 March 2017.

## ROC 曲线和 AUROC 值

ROC 曲线 (Receiver Operating Characteristic curve) 源于信号检测理论，最早由二战中的电子工程师和雷达工程师发明，用来侦测战场上的敌军载具，之后很快就被引入了心理学来进行信号的知觉检测。此后被引入机器学习领域，用来评判分类和检测结果的好坏，是一种常见的统计分析方法，也常被应用于信用评级方法的检验[3]。

下面我们将用图 1 来解释这个指标的原理。图 1 展示了违约的受评对象和未违约的受评对象的信用等级分布，可以看出，由于未违约的受评对象的信用等级更高（信用质量更优），使得信用评级方法具有区分力或排序能力。V 作为截断点 (cut-off point)，提供了一种简单的判定规则来将受评对象区分为未来违约或未违约。该规则下，所有信用等级低于 V 的对象将被视为违约者，所有信用等级高于 V 的对象将被视为未违约者。

**图 1：违约和未违约的受评对象在信用等级下的分布**




*数据来源：鹏元国际*

在这种决策规则下，会有四种情况发生，据此我们构建出了表 1 所示的混淆矩阵(Confusion Matrix)。如果信用等级低于 V 的对象最终真的违约，说明评级方法正确预测了违约对象（击中），我们把被正确击中的对象占所有实际违约的比例为“真阳性率”(True Positive Rate, TPR=Ture Positive / (True Positive +False Negative))；如果信用等级低于 V 的对象最终没有违约，说明评级方法预计错误（错误预警），我们把被错误预警为违约的对象占所有实际未违约的比例为“假阳性率” (False Positive Rate, FPR = False Positive / (False Positive + True Negative))；如果信用等级高于 V 的对象最终违约，说明评级方法预测错误（错误），漏掉了违约对象；如果信用等级高于 V 的对象最终未违约，说明评级方法预测正确（正确预测）。

**表 1：混肴矩阵（根据截断点分类后的四种结果）**

| | | 真实结果 | |
|---|---|---|---|
| | | 违约 | 未违约 |
| 根据评级方法（信用等级）的预测结果 | 违约（截断点及以下） | 击中（真阳性, True Positive, TP） | 错误预警（假阳性, False Positive, FP） |
| | 未违约（截断点以上） | 错误（假阴性, False Negative, FN） | 正确预测（真阴性, True Negative, TN） |

*资料来源：鹏元国际*

每个给定的截断点 V 会对应一组 TPR 和 FPR，假设一共有 k 个可能的 V 值，我们可以得出 k 组 TPR 和 FPR。我们以 FPR 为横坐标的值，TPR 为纵坐标的值画点连线，制成曲线图，即为信用评级对应的 ROC 曲线（见图 2）。可以看出 ROC 曲线开始于

[3] Sobehart JR, Keenan SC (2001), Measuring Default Accurately, Risk 14, pp. S31–S33.

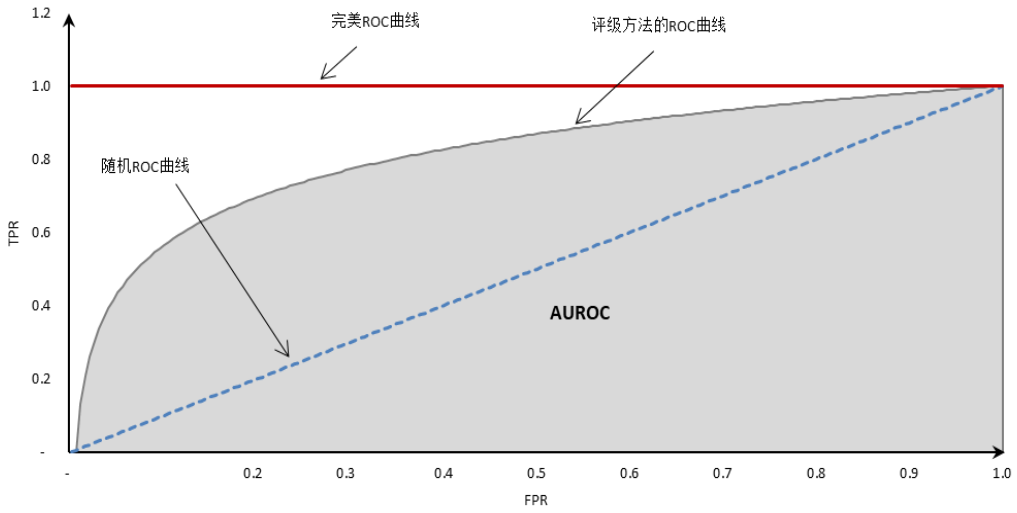
点 (0,0)，终止于点 (1,1) 。如果截断点 V 低于所有可能的信用等级，那么 TPR 和 FPR 均为 0；如果截断点 V 高于所有可能的信用等级，那么 TPR 和 FPR 均为 1。

ROC 曲线有两种极端情况—随机 ROC 曲线和完美 ROC 曲线（见图 2）。对于随机评级方法，其判断违约的规则是随机的，即没有任何区分能力，所以它的 TPR 和 FPR 在所有截断点 V 上都是相等的；在完美排序的信用评级下，图 1 中违约和未违约的信用等级分布是完全不重合的，因此可以完美区分违约和未违约对象。这种情况下 TPR 小于 1 对应的 FPR 全为 0，FPR 大于 0 对应的 TPR 全为 1，所以完美 ROC 曲线是三个点 (0,0), (0,1) 和 (1,1) 连接的直线。

ROC 曲线的信息可以用 AUROC 值来概括，它代表的是 ROC 曲线下的面积。AUROC 取值范围为 0 到 1 之间，其中随机 ROC 曲线的 AUROC 是 0.5，完美 ROC 曲线的 AUROC 是 1.0。AUROC 的值越接近 1，即 ROC 曲线越偏向左上方，信用评级的信用排序能力越强。ROC 曲线的构建和 AUROC 值的详细计算步骤将在附录中给出。

ROC 曲线简单、直观的展示了信用评级的排序能力，是评估信用排序能力的代表性指标。ROC 曲线还有两个很好的特性，首先，ROC 曲线不会随着测试样本（违约/未违约）的分布变化而产生明显变化；另一方面，在实际的数据集中经常会出现样本类不平衡，即正负样本比例差距较大的情况，同时测试数据中的正负样本也可能随着时间变化， ROC 曲线对这些变化不敏感，仍能保持一定的稳定性。

**图 2：CAP 曲线和 AR 值图解**




*数据来源：鹏元国际*

计算出信用评级对应的 AUROC 值以后，我们还可以根据 AUROC 值的一般评价标准（见表 2），对评级的信用排序能力做出相应的评价。在所评样本一致的情况下，我们也可以对不同评级方法得出的 AUROC 值进行对比，从而对它们的信用排序能力进行排序。

**表 2：AUROC 值的一般评价标准**

| AUROC 值 | 评价 |
|---|---|
| [0.9,1.0) | 优秀 |
| [0.8,0.9) | 良好 |
| [0.7,0.8) | 一般 |
| [0.6,0.7) | 较差 |
| [0.5,0.6) | 无效 |

*来源：鹏元国际*

## CAP 曲线和 AR 值

CAP (Cumulative Accuracy Profile) 曲线以及对应的 AR 值是金融风控模型评价中的一个常用指标，其衡量的是风控模型检出风险（违约发行人）的能力，我们同样可以用它来评估信用评级的信用排序能力。

具体而言，以所有受评对象的经验累积分布 $C_T$ 为横轴，违约对象的经验累积分布 $C_D$ 为纵轴，我们可以构造 CAP 曲线（见图 3）。对任意一个给定的信用等级 $R_i$，信用等级低于或等于 $R_i$ 的受评对象数量占所评对象总数的比例记为 $C_T(R_i)$，信用等级低于或等于 $R_i$ 的违约数量占违约对象总数的比例记为 $C_D(R_i)$。假设一共有 k 个信用等级，我们可以计算出 k 组 $C_T(R)$和 $C_T(R)$，最后我们以 $C_T(R_1)$, $C_T(R_2)$, …, $C_T(R_k)$为横坐标的值，$C_D(R_1)$, $C_D(R_2)$, …, $C_D(R_k)$为纵坐标的值画点连线，制成曲线图，即为 CAP 曲线。可以看出，CAP 曲线开始于点 (0,0)，终止于点 (1,1)。

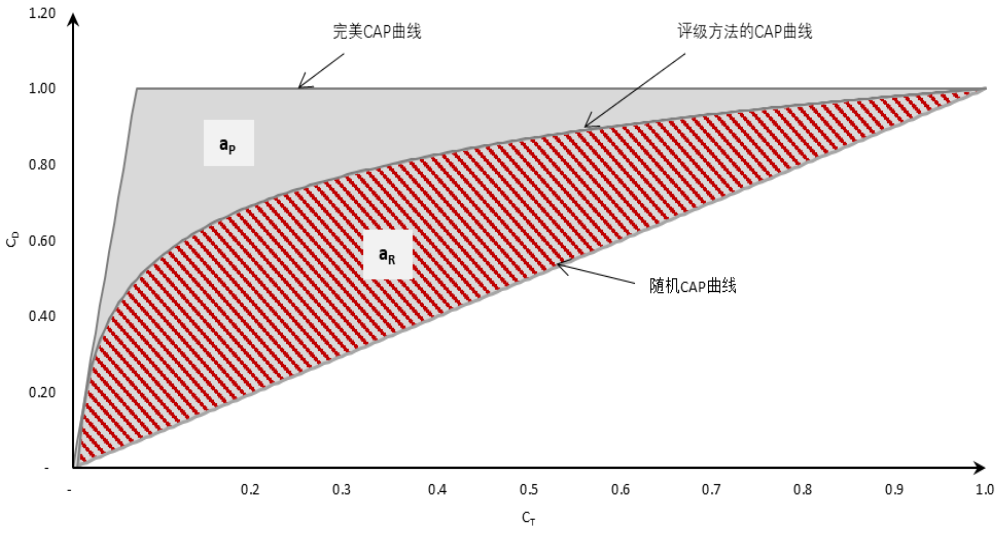
CAP 曲线有两种极端情况，即随机 CAP 曲线和完美 CAP 曲线（见图 3）。对于随机评级方法，其判断违约的规则是随机的，其评级没有任何区分能力，在这种情况下，CAP 曲线是一条平分第一象限的 45 度的直线，因为如果评级方法不包含信用质量信息，它将把 x%的受评对象中 x%分配为违约。完美的评级方法包含完整的信用质量信息，在这种情况下，所有违约的对象都将得到比未违约对象更低的评级，由此产生的 CAP 曲线将直接上升到 1 并一直保持，斜率为 1/违约率。

CAP 曲线中包含的信息可以概括为一个数值，即 AR，计算公式为

$$AR = \frac{a_R}{a_P},$$

其中$a_P$为完美 CAP 曲线和随机 CAP 曲线之间的面积（图 2 中灰色区域），$a_R$的值为评级对应的 CAP 曲线和随机 CAP 曲线之间的面积（图 2 中红色阴影区域）。AR 值的取值范围在 0 到 1 之间，AR 值越接近 1（CAP 曲线越偏向左上方），说明评级的信用排序能力越强。CAP 曲线的构建和 AR 值的详细计算步骤将展示在附录中。目前国际评级机构标普使用了 CAP 曲线对评级质量进行检验，并把 AR 值（也称为基尼系数）作为一项重要指标在每年的全球企业违约研究报告中披露[4]。同时，穆迪也使用了该指标[5]，但是在 2011 年之后的报告中穆迪已经不再发布该指标的检验结果，而由平均违约位置（Average Defaulter Position, ADP）[6] 替代。

**图 3：CAP 曲线和 AR 值图解**




*数据来源：鹏元国际*

需要补充的是，AR 值和 AUROC 值在算法上有相似之处。实质上，这两种度量指标反应的是相同的信用排序能力信息，它们的关系可以用以下公式描述

$$AR = 2 \cdot AUROC - 1$$

[4] “2019 Annual Global Corporate Default and Rating Transition Study” S&P, April,29 2020.

[5] “Special Comment: Measuring: The Performance of Corporate Bond Ratings” Moody, April 2003.

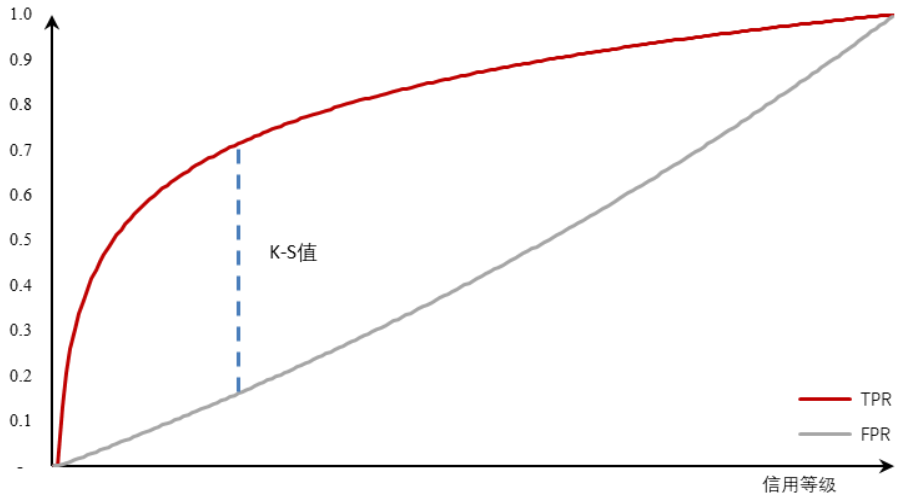
[6] 《中外评级质量对比系列报告之——平均违约位置》中证鹏元, 2019 年 4 月.

### K-S 统计量

K-S 是信用排序能力评估的另一个常用指标，其衡量的是违约和未违约样本累计分布之间的差值。K-S 曲线与 ROC 曲线的构造方法非常类似，也是基于混淆矩阵进行计算，是用另一种方式呈现了信用评级的信用排序能力。不同之处在于，由于 K-S 取的是 TPR 和 FPR 差值的最大值，它能够找到一个最优的截断点；而 AUROC 只评价了信用评级的整体排序能力，并没有指出如何划分类能让排序效果达到最好（没有找到最优的截断点）。

和 ROC 曲线的构造方法类似，我们先构建出表 1 所示的混淆矩阵，然后计算各个截断点的 TPR 和 FPR。假设一共有 k 个可能的 V 值，我们于是可以得出 k 组 TPR 和 FPR，我们以各个截断点为横坐标的值，FPR 和 TPR 为纵坐标的值画点连成两条曲线，制成 K-S 图（见图 4）。随着信用等级逐渐变大，TPR 越快提升，评级的信用排序能力越强；反之，FPR 越快提升，信用排序能力就越差。K-S 值，正是图中的最大差值，此时的横轴取值，便是最优截断点。K-S 值的计算步骤将展示在附录中。

**图 4：K-S 图和 K-S 值图解**




*数据来源：鹏元国际*

计算出 K-S 值以后，我们可以根据 K-S 值的一般评价标准（见表 3），对评级的信用排序能力做出相应的评价。在所评样本一致的情况下，我们也可以对不同的评级方法的 K-S 值进行对比，从而对它们的的信用排序能力进行排序。

**表 3：K-S 值的一般评价标准**

| K-S 值 | 评价 |
|---|---|
| ≥0.75 | 异常 |
| [0.60,0.75) | 优秀 |
| [0.50,0.60) | 良好 |
| [0.40,0.50) | 一般 |
| [0.20,0.40) | 较差 |
| <0.2 | 无效 |

*来源：鹏元国际*

## 预测准确性

预测准确性 (Predictive Accuracy) 衡量的是评级对各信用等级对应的违约率预测的准确性，一般通过观测评级方法预测出的违约率和预期/理想违约率之间是否一致来进行评价。所以，进行预测准确性评估的前提是我们需要获得每个信用等级对应的预期违约率信息（绝对数字或范围），这往往需要使用该评级方法的信用评级机构提供。

对信用评级的预测准确性评估一般使用统计学中的假设检验来完成。常用的适用于单周期检验方法（即从首次引入信用等级至一年后这一时间段做假设检验，并在随后每年开展该检验）有二项式检验(binomial test)和卡方检验 (Chi-square test/Hosmer-

Lemeshow test)；适用于多周期的检验方法（同时对经过多个时间段/几年的评级方法整体做假设检验）的常见检验有正态分布检验 (normal test)。上述评估方法被广泛推荐和应用于验证《巴塞尔协议 II》"内部评级"(IRB)系统中的 PD (probability of default)、LGD (Loss Given Default)、EAD (Exposure at Default) 等参数的准确性和可靠性[7]。下面我们将逐一介绍它们。

### 二项式检验

二项式检验是单周期的假设检验中最常用的方法，该检验评估了一段时间内（一般为一年）信用评级对每一个信用等级下对应的违约率预测的准确性。假设一共有 k 个信用等级，对于每一个信用等级 $R_i$，我们假设在该等级内的所有违约事件是相互独立的。这一假设使得信用等级为 $R_i$ 的违约个数可被看作服从二项式分布的随机数，记作 $X\sim B(N_i,PD_i)$，其中 $N_i$ 为信用等级为 $R_i$ 的受评对象总数，$PD_i$ 为信用等级为 $R_i$ 的预期违约率。因此，我们可以通过检验原假设来评估一段时间内违约率预测的准确性：

H0：评级方法估计出的违约率足够保守，即信用等级 $R_i$ 下实际观测到的违约率 ≤预期违约率 $PD_i$

H1：评级方法低估了违约风险，即信用等级 $R_i$ 下实际观测到的违约率 > 预期违约率 $PD_i$

当在信用等级 $R_i$ 中观察到的违约的数量 $d_i$ 大于或等于临界值 $d_i(\alpha)$时，我们在显著水平 α 下拒绝原假设 H0。临界值 $d_i(\alpha)$的计算公式为

$$d_i(\alpha) = \min\left\{d: \sum_{j=d}^{N_i}\binom{N_i}{j} PD_i^j\left(1-PD_i^j\right)^{N_i-j} \le 1-\alpha\right\}$$

在固定置信水平下，二项式检验在所有检验中是最强大的。但是二项式检验有一个假设前提，即违约事件之间是相互独立的。事实上，该假设是不太实际的，实证中显示大多违约事件之前都是存在相关性的。若违约事件之间相互关联，二项式检验会放大实际违约率和预测违约率之间的差距，即当信用评级低估了违约率时（实际违约率大于预期违约率时），二项式检验会放大这种低估程度。因此，从纯保守的风险评估角度来看，在违约风险低估的情况下，放大低估的程度只会使二项式检验方法更加保守。所以，我们可以在违约事件相互独立的假设下进行二项式检验。二项式检验详细步骤将展示在附录中。

### 卡方检验

卡方检验是单周期假设检验的另一种方法。与二项式检验不同的是，卡方检验提供了一种同时检验多个信用等级下违约率预测准确性的方法。信用等级为 $R_i$ 对于的预期违约率记作 $PD_i$，卡方检验的前提假设为 1) 实际的违约分布和预期的违约分布是一致的；2) 每个信用等级内以及所有信用等级之间的所有违约事件都相互独立。

卡方检验的统计量计算公式如下

$$S_k = \sum_{i=1}^{k}\frac{(N_i\cdot PD_i - d_i)^2}{N_i\cdot PD_i\cdot(1-PD_i)},$$

其中 $N_i$ 代表信用等级 $R_i$ 的所评对象个数，$d_i$ 代表信用等级 $R_i$ 的违约个数。根据中心极限定理可得出，当所有 $N_i$ (i=1, 2, …, k) 趋于无穷时，$S_k$ 将依分布收敛于自由度为 k 的卡方分布。因此，我们可以通过检验原假设来评估一段时间内违约率预测的准确性：

H0：评级方法合理估计了违约率，即对于所有 i (i=1, 2, …, k)，信用等级 $R_i$ 下观测到的违约率=预期违约率 $PD_i$

H1：评级方法没有合理估计违约风险，即存在 i，使得信用等级 $R_i$ 下观测到的违约率 ≠预期违约率 $PD_i$

然后我们便可以通过计算统计量 $S_k$ 对应的 p 值（即自由度为 k 的卡方分布大于 $S_k$ 的概率）来进行检验。在显著水平 α 下，若 p 值小于 α，我们拒绝原假设 H0，反之我们无法拒绝原假设，说明评级方法足够保守（通过检验）。此外，我们还可以依据 p 值直接比较不同评级方法的预测结果，p 值越小说明评级方法的预测准确性表现越差。卡方检验的具体步骤将展示在附录中。

[7] "Studies on the Validation of Internal Rating Systems", Basel Committee on Banking Supervision, May, 2005.

### 正态检验

由于二项式检验和卡方检验仅适用于单周期检验框架，为克服单周期的独立性假设，我们下面将介绍一种适用于多周期（周期间相互关联）的检验方法——正态检验。正态检验是对单一评级等级的违约概率预测准确性的多周期检验。其假设前提是违约率均值随着时间变化不会有大幅的波动，并且不同年份的违约事件之间也是相互独立的，即对于每个信用等级 $R_i$( i=1, 2, …, k)，各年份的违约率 $p_{i,t}$(t = 1,…,T)是独立同分布的，期望记为 $PD_{i,t}$，方差记为 $\sigma_i$。在这些假设下，根据中心极限定理，我们可以得出当时间 T 趋于无穷时，统计量

$$S_i^N = \frac{\sum_{t=1}^{T}(p_{i,t} - PD_{i,t})}{\sigma_i \cdot \sqrt{T}},$$

会依分布收敛于标准正态分布，由于该统计量的收敛速度很快，即使很小的 T 值也可以达到很好的效果。此外，违约率的方差 $\sigma_i$ 需要进一步估计，该方差的常用估计量为

$$\hat{\sigma}_i^2 = \frac{1}{T-1} \cdot \left[ \sum_{t=1}^{T} \left(p_{i,t} - PD_{i,t}\right)^2 - \frac{1}{T}\left( \sum_{t=1}^{T} \left(p_{i,t} - PD_{i,t}\right) \right)^2 \right]。$$

因此，我们可以通过检验原假设来评估一段时间内违约率预测的合理性：

H0：评级方法估计出的违约率足够保守，即对所有 t (t = 1,2,…,T)，信用等级 $R_i$ 下观测到的违约率≤预期违约率 $PD_{i,t}$

H1：评级方法低估了违约风险，即存在 t，使得信用等级 $R_i$ 下观测到的违约率> 预期违约率 $PD_{i,t}$

然后我们便可以对比统计量和临界值来进行检验。在显著水平 α 下，若统计量大于正态分布的 α-分位数点值 $Z_\alpha$，我们拒绝原假设 H0，反之我们无法拒绝原假设，说明评级方法足够保守（通过检验）。正态检验的具体步骤将展示在附录中。

在正态检验中，不同信用等级的违约事件之间的相关性是允许的。正态检验的能力较好，一般情况下第一类错误（type I error）都小于显著水平。此外，若独立性假设不成立，正态检验的效果仍能保持一定的稳健性。然而，对于短周期（例如 5 年），正态检验的表现不及预期。

## 评级稳定性

评级稳定性 (Rating Stability) 衡量的是评级结果（信用等级分布）的稳定性和违约率的稳定性，主要通过分析信用评级的实际信用等级分布和预期信用等级分布之间的差异来评估。

构造评级迁移矩阵和计算级别迁移率是分析评级调整情况以及衡量评级稳定性的常用方法。国际三大评级机构标普、穆迪和惠誉每年都会发布年度等级迁移矩阵数据来检验评级稳定性。具体来说，我们可以通过观察上调率、下调率、持平率和调整幅度，以及它们的变化趋势来进行分析。

另一个常见的分析指标为群体稳定性指标 (population/system stability index, PSI/SSI)，其衡量的是观测的结果及期望的结果的分布之间的差异。PSI 的计算公式如下：

$$PSI = \sum_{i=1}^{k} (Ac_i - Ex_i)(Ln(Ac_i) - Ln(Ex_i)) = \sum_{i=1}^{k} (Ac_i - Ex_i) Ln(Ac_i / Ex_i)$$

其中 $Ac_i$ 为观测到的信用等级为 $R_i$ 的违约占比，$Ex_i$ 为预期的信用等级为 $R_i$ 的违约占比。同样地，各信用等级下期望违约占比需要评级机构提供。

## 评级质量评估方法的实际应用

我们选取了中国主要的八家信用评级机构，以它们的企业评级为研究对象，进行评级质量的量化评估。具体来说，我们将评估这些评级机构的企业评级在信用排序能力、预测准确性、评级稳定性这三个维度的实际表现。我们以这些评级机构在 2014-2019 年间所评的信用债发行人为研究样本，共计 7,671 个发行人和 33,149 条评级信息，每家评级公司涉及的发行人数量和评级信息量如表 4 所示。由于远东资信评估有限公司所评的信用债发行主体的总样本不足，且所评主体均未出现违约，无法满足评估检验要求，我们将不对其进行评估，所以我们的研究对象最终确定为剩余的七家评级机构的企业评级方法。

**表 4：评级机构列表**

| 编号 | 评级机构简称 | 评级机构全称 | | 发行人样本总数 | 历史评级信息笔数 |
|---|---|---|---|---|---|
| 1 | 中诚信 | 中诚信国际信用评级有限责任公司 | 中诚信证券评估有限公司 | 2051 | 7226 |
| 2 | 中债 | 中债资信评估有限责任公司 | | 2575 | 5058 |
| 3 | 联合 | 联合资信评估有限公司 | 联合信用评级有限公司 | 1787 | 6148 |
| 4 | 中证鹏元 | 中证鹏元资信评估股份有限公司 | | 1299 | 4757 |
| 5 | 东方金诚 | 东方金诚国际信用评估有限公司 | | 652 | 1698 |
| 6 | 大公国际 | 大公国际资信评估有限公司 | | 1088 | 4097 |
| 7 | 上海新世纪 | 上海新世纪资信评估投资服务有限公司 | | 1168 | 4130 |
| 8 | 远东 | 远东资信评估有限公司 | | 14 | 35 |

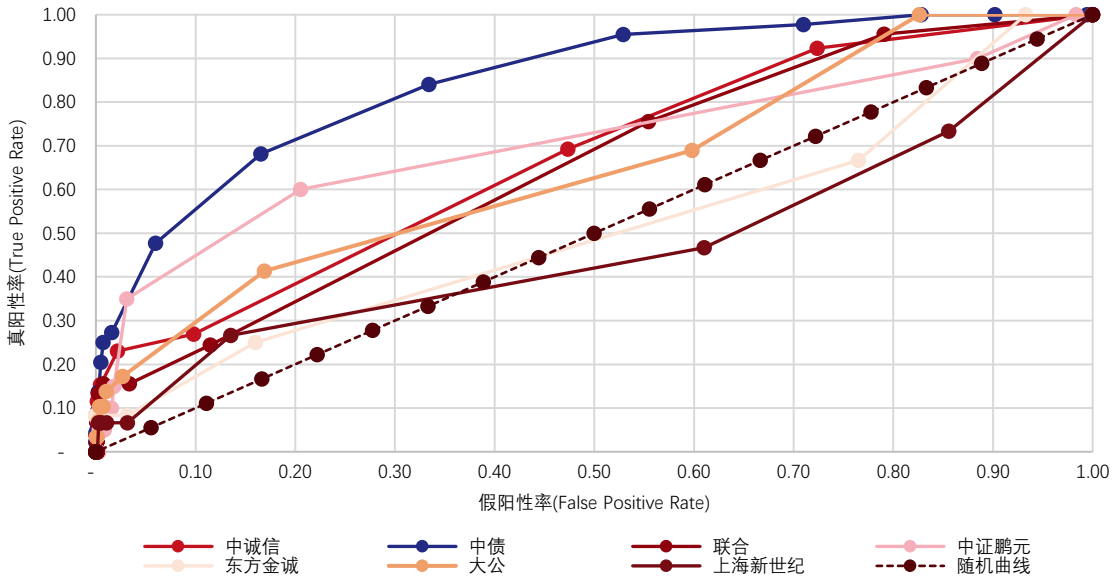
*备注：1.我们把中诚信国际信用评级和中诚信证券评估合并为一家评级机构；同时我们把联合资信和联合信用评级合并为一家评级机构；*
*2. 中债资信的评级信息包括主动评级和委托评级；*
*3. 由于资产支持证券 (ABS)结构上的特殊性，我们研究的评级样本中不包含资产支持证券 (ABS)的信用评级 。*
*来源：鹏元国际，Wind*

### 信用排序能力评估

在信用排序能力评估中，我们以七家评级机构在 2014-2019 年间所评的信用债发行人为研究样本，分析这些评级机构所给出的信用评级能否有效区分未来一年内将违约的发行主体。对于发行人的主体评级数据，我们将采用每年数据采集日（每年的 6 月 30 日）的最新主体信用评级，共计 7,671 个发行人与 33,114 条评级信息，其中 174 个发行人在 2015-2020 期间违约。

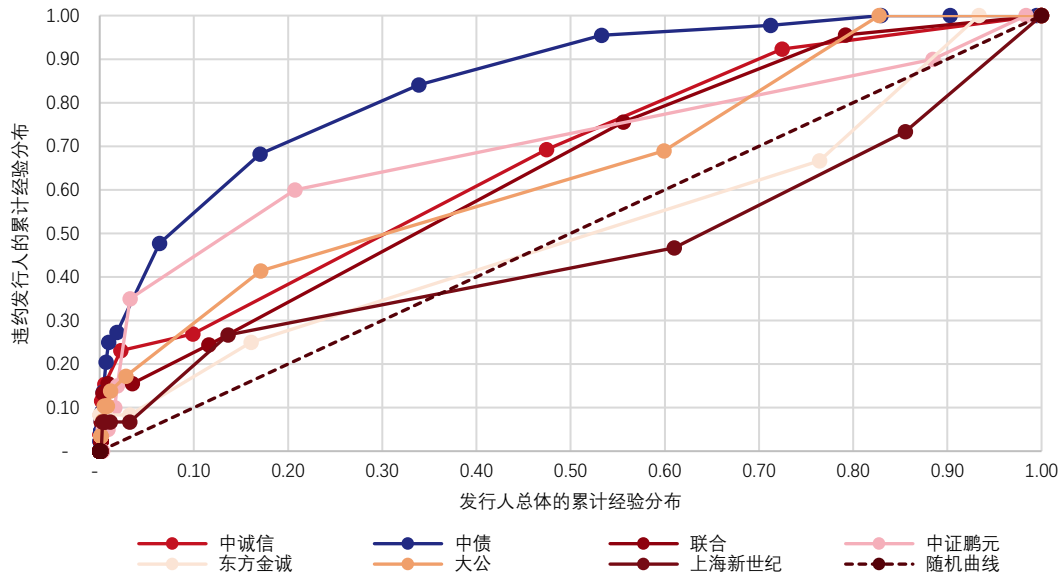
图 5 和图 6 展示了不同评级公司的 ROC 曲线和 CAP 曲线。表 5 展示了根据 ROC 曲线计算出的 AUROC 值以及对应的信用排序能力评价，和根据 CAP 曲线计算出的 AR 值。通过 AUROC 的指标我们可以看出，中债的信用评级的信用排序能力为良好，中证鹏元的信用评级的信用排序能力为一般，中诚信、联合和大公国际评级的信用排序能力较差，上海新世纪和东方金诚评级的信用排序能力是无效的。

**图 5：七家评级公司评级方法的 ROC 曲线**




*数据来源：鹏元国际，Wind，中证鹏元债券评级数据库*

**图 6：七家评级公司评级方法的 CAP 曲线**




*数据来源：鹏元国际，Wind，中证鹏元债券评级数据库*

**表 5：七家评级公司评级方法的 AUROC 值、AR 值以及信用排序能力评价**

| 编号 | 评级机构简称 | AR 值 | AUROC 值 | AUROC 评价 |
|---|---|---|---|---|
| 1 | 中诚信 | 0.34 | 0.67 | 较差 |
| 2 | 中债 | 0.70 | 0.85 | 良好 |
| 3 | 联合 | 0.30 | 0.65 | 较差 |
| 4 | 中证鹏元 | 0.41 | 0.71 | 一般 |
| 5 | 东方金诚 | 0.02 | 0.51 | 无效 |
| 6 | 大公国际 | 0.30 | 0.65 | 较差 |
| 7 | 上海新世纪 | -0.07 | 0.47 | 无效 |

*来源：鹏元国际*

同时，我们计算了衡量评级信用排序能力的第三个指标—K-S 值。表 6 展示了不同评级公司的 K-S 值和相对应的信用排序能力评价。K-S 值的评价结果和 AUROC 值的类似，我们可以看出，中债资信评级的信用排序能力良好，中证鹏元评级的信用排序能力一般，中诚信、联合和大公国际评级的信用排序能力较差，上海新世纪和东方金诚的评级无信用排序能力，和根据 AUROC 值得出的评价一致。

**表 6：七家评级公司评级方法的 K-S 值及其评价**

| 编号 | 评级机构简称 | K-S 值 | K-S 评价 |
|---|---|---|---|
| 1 | 中诚信 | 0.22 | 较差 |
| 2 | 中债 | 0.52 | 良好 |
| 3 | 联合 | 0.20 | 较差 |
| 4 | 中证鹏元 | 0.39 | 一般 |
| 5 | 东方金诚 | 0.10 | 无效 |
| 6 | 大公国际 | 0.24 | 较差 |
| 7 | 上海新世纪 | 0.14 | 无效 |

*来源：鹏元国际*

在本节的信用排序能力评估中，我们考察的是评级机构的信用评级能否有效区分未来一年内将违约的发行人主体。我们同样可以把评级机构给出的信用评级能否有效区分未来两年内或三年内将违约的发行人作为信用排序能力的评判标准，这两种标准下的信用排序能力评估结果将展示在附录中。

### 预测准确性评估

在预测准确性的评估中，我们以七家评级机构在 2014-2019 年间所评的信用债发行人为研究样本，衡量信用评级对下一年期违约率预测的准确性。我们以中国信用债市场整体一年期平均违约率作为预期违约率（见表 7），来衡量各家评级机构信用等级下的一年期违约率是否符合预期（是否存在低估）。我们可以看到在中国信用债市场，发行人主体评级几乎全部集中在等级 AA-以上，因此在本节我们评估评级的预测准确性时，仅衡量信用评级在 AA-以上信用等级的违约率预测的准确性。需要说明的是，为了信用等级的一致性和可比性，我们把中债资信的 AAA+和 AAA-等级划分为 AAA 等级。

**表 7：一年期平均违约率（2014 年-2019 年）**

| 信用等级 | 发行人占比 | 预期违约率 |
|---|---|---|
| **AAA** | 14.54% | 0.1717% |
| **AA+** | 19.43% | 0.4497% |
| **AA** | 46.33% | 0.4131% |
| **AA-** | 12.60% | 0.6935% |
| A+ | 3.38% | 0.7380% |
| A | 1.44% | 1.7291% |
| A- | 0.87% | 0.0000% |
| BBB+ | 0.45% | 3.7383% |
| BBB | 0.21% | 6.0000% |
| BBB- | 0.06% | 6.6667% |
| BB+ | 0.08% | 0.0000% |
| BB | 0.08% | 21.0526% |
| BB- | 0.01% | 0.0000% |
| B+ | 0.01% | 0.0000% |
| B | 0.04% | 0.0000% |
| B- | 0.01% | 0.0000% |
| CCC | 0.02% | 20.0000% |
| CC | 0.05% | 8.3333% |
| C | 0.02% | 0.0000% |

*来源：鹏元国际，Wind，中证鹏元债券评级数据库*

我们首先使用二项式检验方法进行预测准确性评估，二项式检验评估了 2014 年-2019 年各家评级机构的信用评级在各个信用等级下预期违约率估计的准确性，假设检验中的显著性水平设定为 5%。表 8-11 为检验结果，其中 “拒绝” 假设检验代表的是信用评级没有合理预测信用等级所对应的违约率，存在低估违约率情况。可以看出对于信用等级 AAA 而言，预测准确性较好的评级机构有中债、中证鹏元、东方金诚和大公国际；对于信用等级 AA+，预测准确性较好的评级机构有中债和中证鹏元；对于信用等级 AA，预测准确性较好的评级机构有中债、东方金诚和上海新世纪；对于信用等级 AA-，预测准确性较好的评级机构有中诚信和中债。

**表 8：七家评级公司评级方法在“AAA”等级的违约预测准确性二项式检验（2014 年-2019 年）**

| 年份 | 中诚信 | 中债 | 联合 | 中证鹏元 | 东方金诚 | 大公国际 | 上海新世纪 |
|---|---|---|---|---|---|---|---|
| 2014 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 |
| 2015 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 |
| 2016 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 |
| 2017 | 拒绝 | 接受 | 拒绝 | 接受 | 接受 | 接受 | 接受 |
| 2018 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 拒绝 |
| 2019 | 拒绝 | 接受 | 拒绝 | 接受 | 接受 | 接受 | 拒绝 |
| 拒绝年数 | 2 | 0 | 2 | 0 | 0 | 0 | 2 |

*来源：鹏元国际*

**表 9：七家评级公司评级方法在“AA+”等级的违约预测准确性二项式检验（2014 年-2019 年）**

| 年份 | 中诚信 | 中债 | 联合 | 中证鹏元 | 东方金诚 | 大公国际 | 上海新世纪 |
|---|---|---|---|---|---|---|---|
| 2014 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 |
| 2015 | 拒绝 | 接受 | 拒绝 | 接受 | 接受 | 接受 | 拒绝 |
| 2016 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 |
| 2017 | 拒绝 | 接受 | 接受 | 接受 | 接受 | 拒绝 | 接受 |
| 2018 | 拒绝 | 接受 | 拒绝 | 接受 | 拒绝 | 拒绝 | 拒绝 |
| 2019 | 拒绝 | 接受 | 拒绝 | 拒绝 | 拒绝 | 拒绝 | 接受 |
| 拒绝年数 | 4 | 0 | 3 | 1 | 2 | 3 | 2 |

*来源：鹏元国际*

**表 10：七家评级公司评级方法在“AA”等级的违约预测准确性二项式检验（2014 年-2019 年）**

| 年份 | 中诚信 | 中债 | 联合 | 中证鹏元 | 东方金诚 | 大公国际 | 上海新世纪 |
|---|---|---|---|---|---|---|---|
| 2014 | 接受 | 接受 | 拒绝 | 接受 | 接受 | 接受 | 接受 |
| 2015 | 接受 | 接受 | 拒绝 | 接受 | 接受 | 接受 | 拒绝 |
| 2016 | 拒绝 | 接受 | 拒绝 | 接受 | 接受 | 拒绝 | 接受 |
| 2017 | 拒绝 | 接受 | 拒绝 | 拒绝 | 接受 | 拒绝 | 接受 |
| 2018 | 拒绝 | 拒绝 | 拒绝 | 拒绝 | 拒绝 | 拒绝 | 拒绝 |
| 2019 | 拒绝 | 接受 | 拒绝 | 拒绝 | 拒绝 | 拒绝 | 接受 |
| 拒绝年数 | 4 | 1 | 6 | 3 | 2 | 4 | 2 |

*来源：鹏元国际*

**表 11：七家评级公司评级方法在“AA-”等级的违约预测准确性二项式检验（2014 年-2019 年）**

| 年份 | 中诚信 | 中债 | 联合 | 中证鹏元 | 东方金诚 | 大公国际 | 上海新世纪 |
|---|---|---|---|---|---|---|---|
| 2014 | 接受 | 接受 | 接受 | 接受 | 接受 | 拒绝 | 接受 |
| 2015 | 拒绝 | 拒绝 | 接受 | 拒绝 | 接受 | 拒绝 | 拒绝 |
| 2016 | 接受 | 接受 | 拒绝 | 拒绝 | 拒绝 | 拒绝 | 拒绝 |
| 2017 | 接受 | 接受 | 拒绝 | 接受 | 接受 | 拒绝 | 接受 |
| 2018 | 接受 | 接受 | 接受 | 拒绝 | 接受 | 拒绝 | 接受 |
| 2019 | 接受 | 接受 | 拒绝 | 接受 | 拒绝 | 接受 | 接受 |
| 拒绝年数 | 1 | 1 | 3 | 3 | 2 | 5 | 2 |

*来源：鹏元国际*

我们进一步使用卡方检验来整体衡量信用评级在 AA-以上的信用等级的违约率估计是否准确。表 12 为显著水平 5%下的卡方检验结果，其中 “拒绝” 代表的是信用评级没有准确预测所在信用等级下的违约率。可以看出，整体表现最好的评级机构是中债、中诚信和中证鹏元。

**表 12：七家评级公司评级方法在“AA-”及以上的违约预测准确性卡方检验（2014 年-2019 年）**

| 评级机构简称 | p 值 | | | | | | 验证结果 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 拒绝年数 |
| 中诚信 | 0.53 | 0.74 | 0.56 | 0.79 | 0.08 | 0.83 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 0 |
| 中债 | 0.86 | 0.94 | 0.85 | 0.88 | 0.68 | 0.48 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 0 |
| 联合 | 0.82 | 0.90 | 0.74 | 0.41 | 0.00 | 0.01 | 接受 | 接受 | 接受 | 接受 | 拒绝 | 拒绝 | 2 |
| 中证鹏元 | 0.56 | 0.09 | 0.63 | 0.65 | 0.87 | 0.16 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 0 |
| 东方金诚 | 0.99 | 0.96 | 0.40 | 0.85 | 0.52 | 0.00 | 接受 | 接受 | 接受 | 接受 | 接受 | 拒绝 | 1 |
| 大公国际 | 0.82 | 0.54 | 0.91 | 0.94 | 0.00 | 0.00 | 接受 | 接受 | 接受 | 接受 | 拒绝 | 拒绝 | 2 |
| 上海新世纪 | 0.68 | 0.53 | 0.61 | 0.56 | 0.09 | 0.00 | 接受 | 接受 | 接受 | 接受 | 接受 | 拒绝 | 1 |

*来源：鹏元国际*

在本节的预测准确性评估中，我们考察的是评级对一年期的违约率预测是否合理。我们同样可以把评级对两年期或三年期的违约预测准确性作为评判标准，这两种标准的评估结果将展示在附录中。

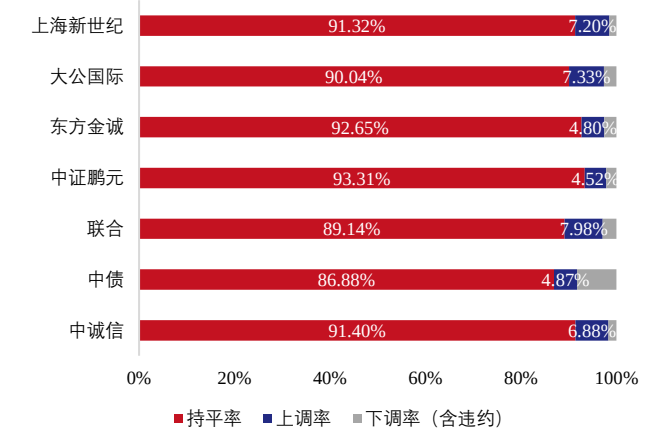
### 评级稳定性评估

我们通过构造评级迁移矩阵和级别迁移率来分析评级调整情况，并以此衡量信用评级在各等级的稳定性。我们使用各家评级机构在 2014-2020 年间的历史评级数据，计算了一至三年期的平均上调率、下调率（含违约）和持平率以及对应调整幅度（见图 7-9）。整体趋势来看，随着统计周期的延长（一年期至三年期），各家评级机构的持平率都在逐渐下降；迁移率方面，除了中债以外，其他六家评级机构的上调率均大于下调率；调整幅度方面，七家评级机构的下调幅度有所差异，但均大于上调幅度，平均上调幅度保持在 1 个级别左右。

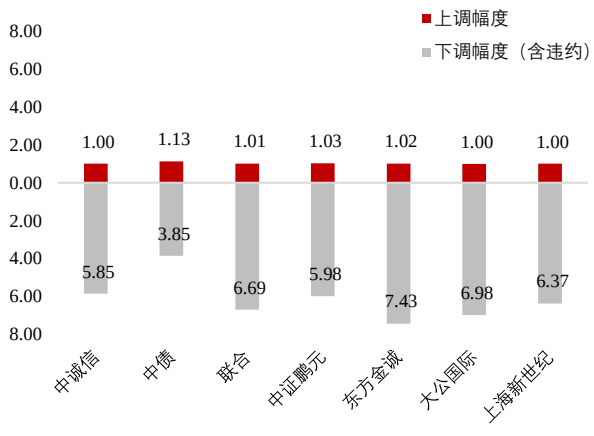
持平率和下调幅度可以从一定程度来衡量信用评级的稳定性和波动幅度。从一年期平均持平率来看，稳定性前三的评级机构为中证鹏元、东方金诚和中诚信；从两年期平均持平率来看，稳定性前三的评级机构为东方金诚、中证鹏元和中诚信；从三年期平均持平率来看，稳定性前三的评级机构为东方金诚、中证鹏元和中债。从一至三年期的平均下调幅度来看，波动幅度较大的评级机构一直为东方金诚、大公国际和联合。综合上述分析，我们可以看出评级稳定性较好的三家评级公司为中证鹏元、中诚信和东方金城。

衡量评级稳定性的另一个常见的分析指标为群体稳定性指标，其衡量的是观测的评级结果及期望的评级结果的分布差异。由于我们无法获取各家评级公司评级方法的预期违约率数值，也没有足够信息对其进行估计，故无法计算该指标，在此节中对此将不再进行展开。

**图 7: 七家评级公司一年期平均调整率（2014-2019 年）**



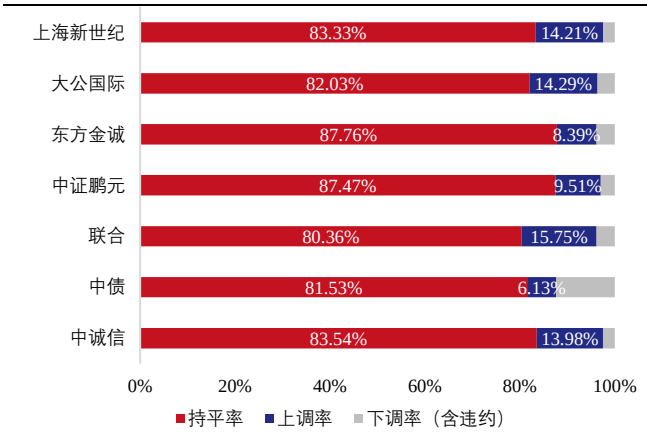

**七家评级公司一年期平均调整幅度（2014-2019 年）**



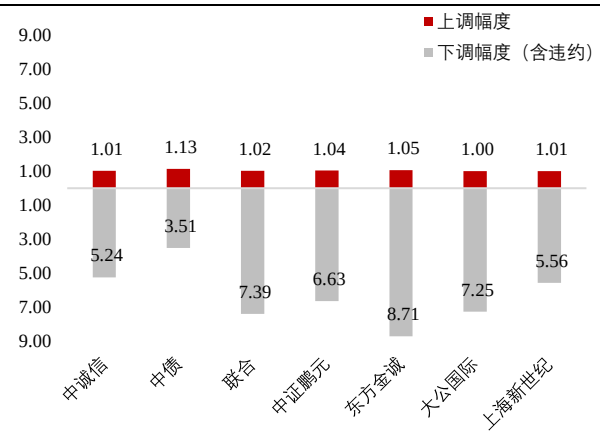

*数据来源：鹏元国际，Wind，中证鹏元债券评级数据库*

*数据来源：鹏元国际，Wind，中证鹏元债券评级数据库*

**图 8: 七家评级公司两年期平均调整率（2014-2018 年）**



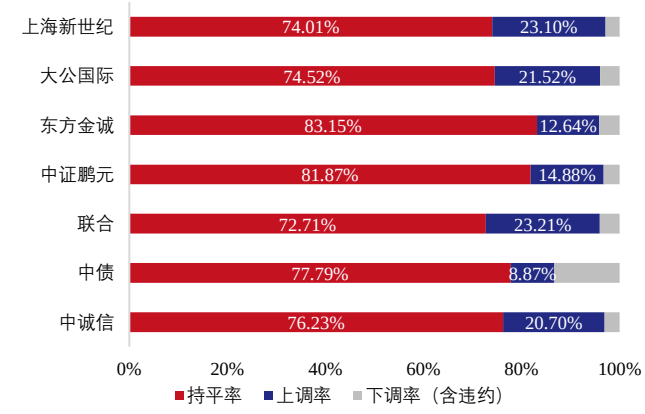

**七家评级公司两年期平均调整幅度（2014-2018 年）**



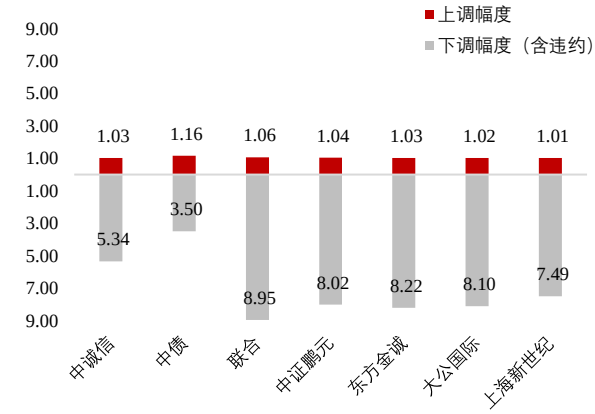

*数据来源：鹏元国际，Wind，中证鹏元债券评级数据库*

*数据来源：鹏元国际，Wind，中证鹏元债券评级数据库*

**图 9: 七家评级公司三年期平均调整率（2014-2017 年）**




**七家评级公司三年期平均调整幅度（2014-2017 年）**




*数据来源：鹏元国际，Wind，中证鹏元债券评级数据库*

*数据来源：鹏元国际，Wind，中证鹏元债券评级数据库*

## 结论

本文借鉴国际先进经验，总结并构建了评级质量的量化评估体系，我们构建的量化评估体系分为三个维度——信用排序能力、预测准确性和评级稳定性。信用排序能力代表的是信用评级区分信用质量好的主体和信用质量差的主体的能力，常用的指标有 AUROC、AR 和 K-S 统计量。预测准确性衡量的是信用评级对各信用等级对应的违约率预测的准确性，一般使用统计学中的假设检验来完成，常用的检验方法有二项式检验，卡方检验和正态分布检验。评级稳定性代表的是评级结果（信用等级分布）的稳定性和违约率的稳定性，常用的评估方法有对评级迁移率的分析和群体稳定性指标(PSI)。

我们进一步的使用我们构建的评级质量评估方法对中国主要七家信用评级机构的企业评级进行了实例展示。评估结果显示，在信用排序能力评估中表现较好的三家机构为中债、中证鹏元和中诚信；整体预测准确性较好的三家机构为中债、中诚信和中证鹏元；评级稳定性较好的三家机构为中证鹏元、中诚信和东方金城。

信用评级机构应采用严格、系统、一致的评级方法，并根据历史经验对其评级质量进行验证（例如回测）。本文通过梳理总结信用评级机构的评级质量评估方法，旨在为中国的信用评级机构、发起人和监管部门等各方提供更丰富有效的评级质量评估工具。我们构建的评估体系一方面可以帮助信用评级机构进行内部检验，加强自律性管理，另一方面可以帮助发起人找寻拥有合理有效评级方法的评级机构，同时还可以为评级业监管部门提供一套全面的信用评级质量评估体系，通过量化指标评价信用评级的有效性，提高监管效率。

# 附录

## ROC 曲线构造方法和 AUROC 计算步骤

1. 确定信用等级集合，记为 **R**={$R_1$,$R_2$,…,$R_k$}，集合中信用等级按信用质量从低到高排列，即 $R_1$ 为最差的信用等级，$R_k$ 为最优的信用等级。一般情况下，我们把 k 设定为 19，同时 **R**= {C, CC, CCC, B-, B, B+, BB-, BB, BB+, BBB-, BBB, BBB+, A-, A, A+, AA-, AA, AA+, AAA}；
2. 选取信用等级集合中的 $R_1$ 作为截断点，并基于 $R_1$ 对每个所评对象做是否违约的预测。预测规则为：若对评对象的信用等级高于截断点，该对象不会违约；反之，该对象将会违约；
3. 计算 $R_1$ 对应的假阳性率(False Positive Rate, FPR) 和真阳性率 (True Positive Rate, TPR)：
   FPR($R_1$) = 预测违约但实际未违约的所评对象的数量 / 总未违约数量；
   TPR($R_1$) = 预测违约且实际违约的所评对象的数量 / 总违约数量；
4. 分别选取信用等级集合中的 $R_2$,…,$R_k$ 作为截断点，重复步骤 2-3，计算对应的假阳性率和真阳性率—TPR($R_i$)和 FPR($R_i$), i=2, 3, …, k；
5. 以 FPR($R_1$), FPR($R_2$), …, FPR($R_k$) 为横坐标的值，TPR($R_1$), TPR($R_2$), …, TPR($R_k$) 为纵坐标的值画点连线，制成曲线图，即为评级方法对应的 ROC 曲线；
   * 在理想状态下，完美的评级方法在任何截断点上的 TPR 和 FPR 应该为 1 和 0，于是我们可以得出完美的 ROC 曲线为一条截距为 1 的水平线；而对于随机的评级方法，在任何截断点上的 TPR 和 FPR 都应相等，所以其 ROC 曲线为一条斜率为 1 的直线。
6. 计算评级方法的 ROC 曲线下面积，即为 AUROC 的值。

## CAP 曲线构造方法和 AR 计算步骤

1. 确定信用等级集合，记为 **R**={$R_1$,$R_2$,…,$R_k$}，集合中信用等级按信用质量从低到高排列，即 $R_1$ 为最差的信用等级，$R_k$ 为最优的信用等级。一般情况下，我们把 k 设定为 19，同时 **R**= {C, CC, CCC, B-, B, B+, BB-, BB, BB+, BBB-, BBB, BBB+, A-, A, A+, AA-, AA, AA+, AAA}；
2. 构造所评对象的累计经验分布，记为 $C_T(R_i)$：
3. $C_T(R_i)$ = 信用等级低于或等于 $R_i$ 的总数/所评对象总数， i=2, 3, …, k；
4. 构造违约对象的累计经验分布，记为 $C_D(R_i)$：
5. $C_D$(Ri) = 信用等级低于或等于 $R_i$ 的违约总数/违约总数, i=2, 3, …, k；
6. 以 $C_T$ ($R_1$), $C_T$ ($R_2$), …, $C_T$ ($R_k$)为横坐标的值，$C_D$ ($R_1$), $C_D$ ($R_2$), …, $C_D$ ($R_k$)为纵坐标的值画点连线，制成曲线图，即为 CAP 曲线；
7. 计算随机 CAP 曲线的面积，为 0.5。随机 CAP 曲线为斜率为 1 的一条直线；
8. 计算评级方法对应的 CAP 曲线和随机 CAP 曲线之间的面积，即为 $a_R$ 的值；
9. 计算完美 CAP 曲线和随机 CAP 曲线之间的面积，即为 $a_P$ 的值。完美 CAP 曲线为开始时斜率为 1/违约率, 上升至 1 后变平的折线，所以我们可以得出 $a_P$=0.5*(1-违约率)；
10. 计算 AR = $a_R/a_P$。

## K-S 统计量计算步骤

1. 确定信用等级集合，记为 **R**={$R_1$,$R_2$,…,$R_k$}，集合中信用等级按信用质量从低到高排列，即 $R_1$ 为最差的信用等级，$R_k$ 为最优的信用等级。一般情况下，我们把 k 设定为 19，同时 **R**= {C, CC, CCC, B-, B, B+, BB-, BB, BB+, BBB-, BBB, BBB+, A-, A, A+, AA-, AA, AA+, AAA}；
2. 选取信用等级集合中的 $R_1$ 作为截断点（cut-off point），并基于 $R_1$ 对每个所评对象是否违约做出预测。预测规则为：若该对象的信用等级高于截断点，则其不会违约；反之，该对象将会违约；
3. 计算 $R_1$ 下的假阳性率(False Positive Rate, FPR) 和真阳性率 (True Positive Rate, TPR)，并计算 KS($R_1$)：
   FPR($R_1$) = 预测违约但实际未违约的数量 / 总未违约数量；
   TPR($R_1$) = 预测违约且实际违约的数量 / 总违约数量；
   KS($R_1$) = | TPR($R_1$)- FPR($R_1$) | ；

4. 分别选取信用等级集合中的 $R_2,\ldots,R_k$ 作为截断点，重复步骤 2-3，计算各自的 $TPR(R_i)$和 $FPR(R_i)$, 从而得出 $KS(R_i)=|TPR(R_i)- FPR(R_i)|$，i=2, 3, …, k;
5. 计算最终 $K\text{-}S = \max\{ KS(R_1), KS(R_2), \ldots, KS(R_k)\}$ 。

### 二项式检验步骤

1. 确定信用等级集合，记为 **R**={$R_1,R_2,\ldots,R_k$}，一般情况下，我们把 k 设定为 19，同时 **R**= {C, CC, CCC, B-, B, B+, BB-, BB, BB+, BBB-, BBB, BBB+, A-, A, A+, AA-, AA, AA+, AAA}；
2. 确定每个信用等级下的预期违约率，记为 **PD**={$PD_1,PD_2,\ldots,PD_k$}；
3. 对每个信用等级 $R_i$ ,i=2, 3, …, k；我们构建以下假设检验：
   $H_0$：信用等级 $R_i$ 下观测到的违约率 ≤预期违约率 $PD_i$；（说明评级方法足够保守）
   $H_1$：信用等级 $R_i$ 下观测到的违约率 > 预期违约率 $PD_i$；（说明评级方法低估了违约风险）
4. 计算每个信用等级 $R_i$ 对应的显著水平 α 下的临界值：

$$d_i(\alpha) = \min\left\{d: \sum_{j=d}^{N_i}\binom{N_i}{j} PD_i^j\left(1-PD_i^j\right)^{N_i-j} \le 1-\alpha\right\},$$

   其中 $N_i$ 为信用等级为 $R_i$ 的所评对象个数；

5. 计算信用等级为 $R_i$ 的违约个数，记为 $d_i$。如果 $d_i > d_i(\alpha)$，我们拒绝原假设 $H_0$，说明评级方法在信用等级 $R_i$ 下的违约率存在低估；反之，说明评级方法足够保守。

### 卡方检验步骤

1. 确定信用等级集合，记为 **R**={$R_1,R_2,\ldots,R_k$}，一般情况下，我们把 k 设定为 19，同时 **R**= {C, CC, CCC, B-, B, B+, BB-, BB, BB+, BBB-, BBB, BBB+, A-, A, A+, AA-, AA, AA+, AAA}；
2. 确定每个信用等级下的预期违约率，记为 **PD**={$PD_1,PD_2,\ldots,PD_k$}；
3. 对多个信用等级，我们构建以下假设检验：
   $H_0$：对于所有 i (i=1, 2, …, k)，信用等级 $R_i$ 下观测到的违约率≤预期违约率 $PD_i$；（说明评级方法足够保守）
   $H_1$：存在 i，使得信用等级 $R_i$ 下观测到的违约率 > 预期违约率 $PD_i$；（说明评级方法低估了违约风险）
4. 计算统计量：

$$S_k = \sum_{i=1}^{k}\frac{(N_i\cdot PD_i - d_i)^2}{N_i\cdot PD_i\cdot(1-PD_i)},$$

   其中 $N_i$ 代表信用等级 $R_i$ 的所评对象个数，$d_i$ 代表信用等级 $R_i$ 的违约个数；

5. 显著水平 α 下，如果 $S_k > \chi_\alpha^2(k)$，我们拒绝原假设 $H_0$，说明评级方法的整体违约率预测存在低估；反之，说明评级方法足够保守。$\chi_\alpha^2(k)$为自由度 k 的卡方分布的 α-分位数点。

### 正态检验步骤

1. 确定信用等级集合，记为 **R**={$R_1,R_2,\ldots,R_k$}，一般情况下，我们把 k 设定为 19，同时 **R**= {C, CC, CCC, B-, B, B+, BB-, BB, BB+, BBB-, BBB, BBB+, A-, A, A+, AA-, AA, AA+, AAA}；
2. 确定不同年份 t (t = 1,…,T), 不同信用等级 $R_i$ (i=1, 2, …, k) 下的预期违约率，记为 $PD_{i,t}$;
3. 计算不同年份 t (t = 1,…,T), 不同信用等级 $R_i$ (i=1, 2, …, k) 下观测到的违约率，记为 $p_{i,t}$;
4. 对每个信用等级 $R_i$ ,i=1, 2, …, k；我们构建以下假设检验：
   $H_0$：对所有 t (t = 1,2,…,T)，信用等级 $R_i$ 下观测到的违约率 $p_{i,t}$ ≤预期违约率 $PD_{i,t}$；（说明评级方法足够保守）
   $H_1$：存在 t，使得信用等级 $R_i$ 下观测到的违约率 $p_{i,t}$ > 预期违约率 $PD_{i,t}$；（说明评级方法低估了违约风险）
5. 计算统计量：

$$S_i^N = \frac{\sum_{t=1}^{T}(p_{i,t}-PD_{i,t})}{\hat{\sigma}_i^2\cdot\sqrt{T}},$$

$$\hat{\sigma}_i^2 = \frac{1}{T-1}\cdot\left[\sum_{t=1}^{T}\left(p_{i,t}-PD_{i,t}\right)^2 - \frac{1}{T}\left(\sum_{t=1}^{T}\left(p_{i,t}-PD_{i,t}\right)\right)^2\right].$$

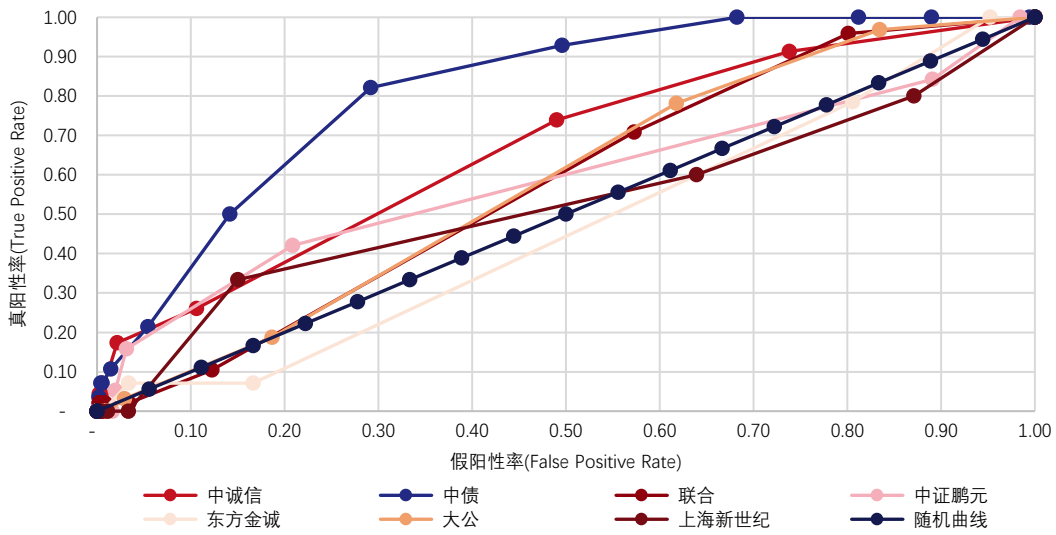
6. 显著水平 α 下，如果统计量 > $Z_\alpha$，我们拒绝原假设 $H_0$，说明评级方法的整体违约率预测存在低估；反之，说明该评级方法足够保守。$Z_\alpha$ 为正态分布的 α-分位数点。

## 七家评级机构的信用评级的信用排序能力评估

正文中在对中国主要七家评级公司的信用评级的信用排序能力进行评估时，我们使用的定义是信用评级是否能有效区分一年内将要违约的发行人。在附录中，我们把定义改为能否有效区分两年/三年内将违约的发行人，重新进行信用排序能力评估。

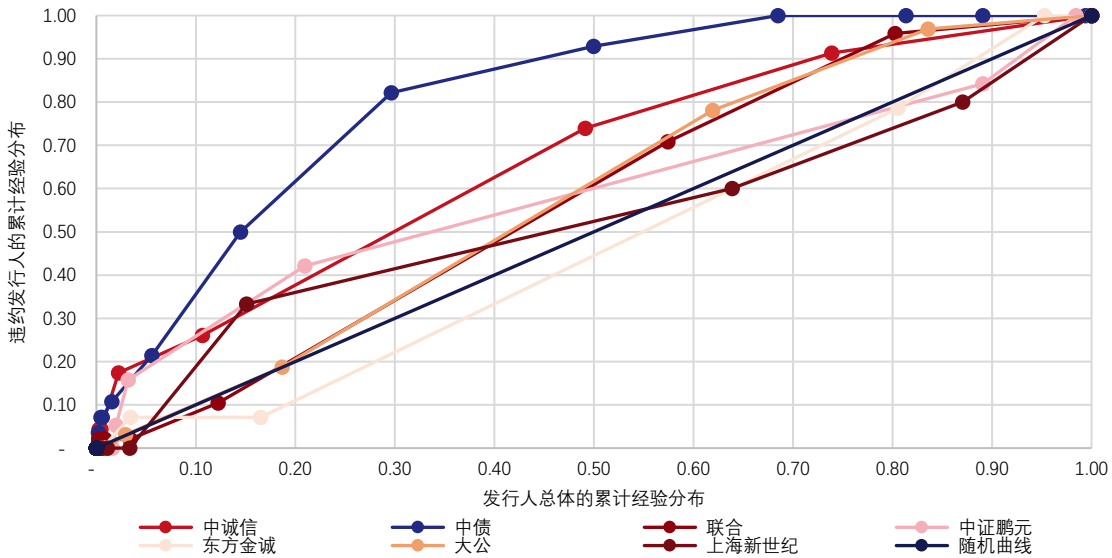
首先我们展示定义改为能否有效区分两年内将违约的发行人后的信用排序能力评估结果。我们计算得出的 ROC 曲线和 CAP 曲线如图 10-11 所示，相对应计算出的 AUROC 值和 AR 值，以及计算出的 K-S 值，以及信用排序能力评价如表 13 所示。

**图 10：七家评级公司评级的 ROC 曲线——有效区分两年内违约**




*数据来源：鹏元国际，Wind，中证鹏元债券评级数据库*

**图 11：七家评级公司评级的 CAP 曲线——有效区分两年内违约**




*数据来源：鹏元国际，Wind，中证鹏元债券评级数据库*

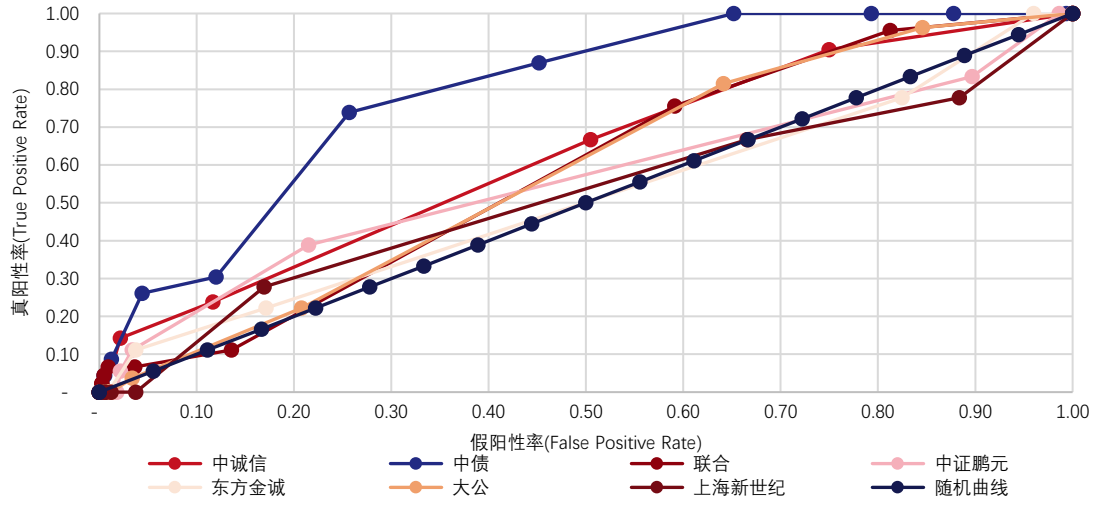
**表 13：七家评级公司评级的 AR 值、AUROC 值以及 K-S 值及其评价——有效区分两年内违约**

| 编号 | 评级机构简称 | AR 值 | AUROC 值 | AUROC 评价 | K-S 值 | K-S 评价 |
|---|---|---|---|---|---|---|
| 1 | 中诚信 | 0.34 | 0.67 | 较差 | 0.25 | 勉强接受 |
| 2 | 中债 | 0.63 | 0.81 | 良好 | 0.53 | 很好 |
| 3 | 联合 | 0.15 | 0.57 | 无效 | 0.16 | 无区别能力 |
| 4 | 中证鹏元 | 0.17 | 0.59 | 无效 | 0.21 | 勉强接受 |
| 5 | 东方金诚 | -0.07 | 0.46 | 无效 | 0.09 | 无区别能力 |
| 6 | 大公国际 | 0.16 | 0.58 | 无效 | 0.16 | 无区别能力 |
| 7 | 上海新世纪 | 0.05 | 0.53 | 无效 | 0.18 | 无区别能力 |

*来源：鹏元国际*

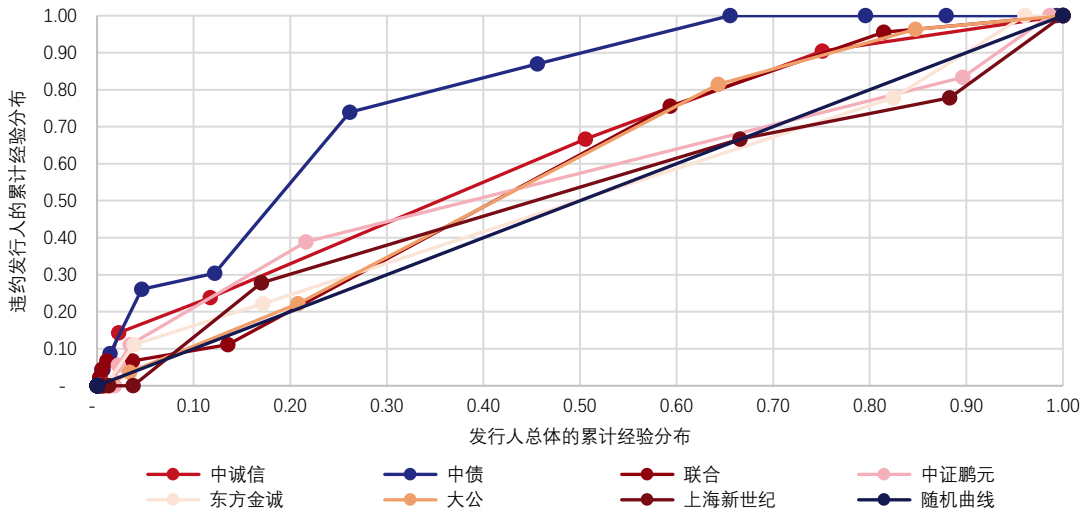
若评级的信用排序能力定义为能否有效区分三年内将违约的发行人，我们得出的 ROC 曲线和 CAP 曲线如图 12-13 所示，相对应计算出的 AUROC 值、AR 值和 K-S 值以及信用排序能力评价如表 14 所示。可以看出各家评级在有效区分三年内讲违约的发行人的能力较差，计算结果显示绝大多数评级机构的评级没有区分能力。这一方面由于评价方法的长期信用排序能力较差，另一方面也有中国评级和违约的历史数据不够充足的原因（中国企业信用债最早违约发生于 2014 年，用于计算信用排序能力的数据长度只有 4 年）。

**图 12：七家评级公司评级的 ROC 曲线——有效区分三年内违约**




*数据来源：鹏元国际，Wind，中证鹏元债券评级数据库*

**图 13：七家评级公司评级的 CAP 曲线——有效区分三年内违约**




*数据来源：鹏元国际，Wind，中证鹏元债券评级数据库*

**表 14：七家评级公司评级的 AR 值、AUROC 值以及 K-S 值及其评价——有效区分三年内违约**

| 编号 | 评级机构简称 | AR 值 | AUROC 值 | AUROC 评价 | K-S 值 | K-S 评价 |
|---|---|---|---|---|---|---|
| 1 | 中诚信 | 0.25 | 0.63 | 较差 | 0.16 | 无区别能力 |
| 2 | 中债 | 0.58 | 0.79 | 一般 | 0.48 | 有区别能力 |
| 3 | 联合 | 0.16 | 0.58 | 无效 | 0.16 | 无区别能力 |
| 4 | 中证鹏元 | 0.12 | 0.56 | 无效 | 0.17 | 无区别能力 |
| 5 | 东方金诚 | 0.02 | 0.51 | 无效 | 0.07 | 无区别能力 |
| 6 | 大公国际 | 0.16 | 0.58 | 无效 | 0.17 | 无区别能力 |
| 7 | 上海新世纪 | 0.03 | 0.51 | 无效 | 0.11 | 无区别能力 |

*来源：鹏元国际*

## 七家评级机构的信用评级方法预测准确性评估

正文中在对中国主要七家评级公司的信用评级的预测准确性进行评估时，我们使用的定义是信用评级对下一年期违约率预测的准确性。在附录中，我们把定义改为信用评级对未来两年/三年违约率预测的准确性，并重新进行评估。

首先，我们衡量七家评级机构的信用评级对未来两年违约率预测的准确性。我们以中国信用债市场整体两年期平均违约率作为预期违约率（见表 15），来衡量各家评级机构的信用评级的两年期违约率是否符合预期（是否存在低估）。同样地，我们仅衡量在 AA-以上信用等级的违约率预测的准确性。表 16-19 为二项式检验结果，表 20 为卡方检验结果，显著性均为 5%。

**表 15：两年期平均违约率（2014 年-2018 年）**

| 信用等级 | 预期违约率 |
|---|---|
| **AAA** | 0.2289% |
| **AA+** | 0.5742% |
| **AA** | 0.6649% |
| **AA-** | 0.6682% |
| A+ | 1.0753% |
| A | 0.7491% |
| A- | 1.7964% |
| BBB+ | 2.4096% |
| BBB | 2.4390% |
| BBB- | 0.0000% |
| BB+ | 7.1429% |
| BB | 0.0000% |
| BB- | 0.0000% |
| B+ | 0.0000% |
| B | 0.0000% |
| B- | 0.0000% |
| CCC | 0.0000% |
| CC | 0.0000% |
| C | 0.0000% |

*来源：鹏元国际，Wind，中证鹏元债券评级数据库*

**表 16：七家评级公司信用评级在“AAA”等级的两年期违约预测准确性二项式检验（2014 年-2018 年）**

| 年份 | 中诚信 | 中债 | 联合 | 中证鹏元 | 东方金诚 | 大公国际 | 上海新世纪 |
|---|---|---|---|---|---|---|---|
| 2014 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 |
| 2015 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 |
| 2016 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 |
| 2017 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 拒绝 |
| 2018 | 拒绝 | 接受 | 拒绝 | 接受 | 接受 | 拒绝 | 拒绝 |
| 拒绝年数 | 1 | 0 | 1 | 0 | 0 | 1 | 2 |

*来源：鹏元国际*

**表 17：七家评级公司信用评级在“AA+”等级的两年期违约预测准确性二项式检验（2014 年-2018 年）**

| 年份 | 中诚信 | 中债 | 联合 | 中证鹏元 | 东方金诚 | 大公国际 | 上海新世纪 |
|---|---|---|---|---|---|---|---|
| 2014 | 拒绝 | 接受 | 拒绝 | 接受 | 接受 | 接受 | 拒绝 |
| 2015 | 接受 | 接受 | 拒绝 | 接受 | 接受 | 接受 | 接受 |
| 2016 | 拒绝 | 接受 | 拒绝 | 接受 | 接受 | 接受 | 接受 |
| 2017 | 拒绝 | 接受 | 拒绝 | 接受 | 拒绝 | 拒绝 | 拒绝 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2018 | 拒绝 | 接受 | 拒绝 | 拒绝 | 拒绝 | 拒绝 | 接受 |
| 拒绝年数 | 4 | 0 | 5 | 3 | 3 | 5 | 3 |

*来源：鹏元国际*

**表 18：七家评级公司信用评级在“AA”等级的两年期违约预测准确性二项式检验（2014 年-2018 年）**

| 年份 | 中诚信 | 中债 | 联合 | 中证鹏元 | 东方金诚 | 大公国际 | 上海新世纪 |
|---|---|---|---|---|---|---|---|
| 2014 | 接受 | 接受 | 拒绝 | 拒绝 | 接受 | 拒绝 | 拒绝 |
| 2015 | 拒绝 | 接受 | 拒绝 | 接受 | 接受 | 拒绝 | 接受 |
| 2016 | 接受 | 接受 | 拒绝 | 拒绝 | 拒绝 | 拒绝 | 拒绝 |
| 2017 | 拒绝 | 接受 | 拒绝 | 拒绝 | 拒绝 | 拒绝 | 拒绝 |
| 2018 | 拒绝 | 接受 | 拒绝 | 接受 | 拒绝 | 拒绝 | 接受 |
| 拒绝年数 | 3 | 0 | 5 | 3 | 3 | 5 | 3 |

*来源：鹏元国际*

**表 19：七家评级公司信用评级在“AA-”等级的两年期违约预测准确性二项式检验（2014 年-2018 年）**

| 年份 | 中诚信 | 中债 | 联合 | 中证鹏元 | 东方金诚 | 大公国际 | 上海新世纪 |
|---|---|---|---|---|---|---|---|
| 2014 | 拒绝 | 拒绝 | 拒绝 | 拒绝 | 接受 | 拒绝 | 拒绝 |
| 2015 | 接受 | 接受 | 接受 | 接受 | 接受 | 拒绝 | 拒绝 |
| 2016 | 接受 | 接受 | 拒绝 | 接受 | 接受 | 接受 | 接受 |
| 2017 | 接受 | 接受 | 拒绝 | 拒绝 | 接受 | 拒绝 | 拒绝 |
| 2018 | 接受 | 拒绝 | 拒绝 | 拒绝 | 接受 | 接受 | 接受 |
| 拒绝年数 | 1 | 2 | 4 | 3 | 0 | 3 | 3 |

*来源：鹏元国际*

**表 20：七家评级公司信用评级在“AA-”及以上的两年期违约预测准确性卡方检验（2014 年-2018 年）**

| 评级机构简称 | p 值 | | | | | 验证结果 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2014 | 2015 | 2016 | 2017 | 2018 | 2014 | 2015 | 2016 | 2017 | 2018 | 拒绝年数 |
| 中诚信 | 0.29 | 0.50 | 0.55 | 0.25 | 0.59 | 接受 | 接受 | 接受 | 接受 | 接受 | 0 |
| 中债 | 0.90 | 0.70 | 0.81 | 0.85 | 0.91 | 接受 | 接受 | 接受 | 接受 | 接受 | 0 |
| 联合 | 0.89 | 0.79 | 0.88 | 0.00 | 0.06 | 接受 | 接受 | 接受 | 拒绝 | 接受 | 1 |
| 中证鹏元 | 0.73 | 0.43 | 0.71 | 0.74 | 0.01 | 接受 | 接受 | 接受 | 接受 | 拒绝 | 1 |
| 东方金诚 | 0.98 | 0.92 | 0.98 | 0.39 | 0.00 | 接受 | 接受 | 接受 | 接受 | 拒绝 | 1 |
| 大公国际 | 0.66 | 0.60 | 0.53 | 0.00 | 0.02 | 接受 | 接受 | 接受 | 拒绝 | 拒绝 | 2 |
| 上海新世纪 | 0.54 | 0.54 | 0.64 | 0.07 | 0.01 | 接受 | 接受 | 接受 | 接受 | 拒绝 | 1 |

*来源：鹏元国际*

使用同样的方法，我们再次衡量七家评级机构的信用评级对未来三年违约率预测的准确性。我们以中国信用债市场整体三年期平均违约率作为预期违约率（见表 21），来衡量各家评级机构信用评级未来三年违约率是否符合预期（是否存在低估）。表 22-25 为二项式检验结果，表 26 为卡方检验结果，显著性均为 5%。

**表 21：三年期平均违约率（2014 年-2017 年）**

| 信用等级 | 预期违约率 |
|---|---|
| **AAA** | 0.3717% |
| **AA+** | 0.6603% |
| **AA** | 0.7019% |
| **AA-** | 0.7689% |
| A+ | 1.2448% |
| A | 2.0202% |
| A- | 0.8621% |
| BBB+ | 6.2500% |
| BBB | 5.7143% |
| BBB- | 0.0000% |
| BB+ | 10.0000% |
| BB | 0.0000% |
| BB- | 0.0000% |
| B+ | 0.0000% |
| B | 0.0000% |
| B- | 0.0000% |
| CCC | 0.0000% |
| CC | 0.0000% |
| C | 0.0000% |

*来源：鹏元国际，Wind，中证鹏元债券评级数据库*

**表 22：七家评级公司信用评级在“AAA”等级的三年期违约预测准确性二项式检验（2014 年-2017 年）**

| 年份 | 中诚信 | 中债 | 联合 | 中证鹏元 | 东方金诚 | 大公国际 | 上海新世纪 |
|---|---|---|---|---|---|---|---|
| 2014 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 |
| 2015 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 |
| 2016 | 接受 | 接受 | 接受 | 接受 | 接受 | 接受 | 拒绝 |
| 2017 | 拒绝 | 接受 | 拒绝 | 接受 | 接受 | 拒绝 | 拒绝 |
| 拒绝年数 | 1 | 0 | 1 | 0 | 0 | 1 | 2 |

*来源：鹏元国际*

**表 23：七家评级公司信用评级在“AA+”等级的三年期违约预测准确性二项式检验（2014 年-2017 年）**

| 年份 | 中诚信 | 中债 | 联合 | 中证鹏元 | 东方金诚 | 大公国际 | 上海新世纪 |
|---|---|---|---|---|---|---|---|
| 2014 | 拒绝 | 接受 | 拒绝 | 接受 | 接受 | 接受 | 拒绝 |
| 2015 | 接受 | 接受 | 拒绝 | 接受 | 接受 | 接受 | 接受 |
| 2016 | 拒绝 | 接受 | 拒绝 | 接受 | 接受 | 拒绝 | 拒绝 |
| 2017 | 拒绝 | 接受 | 拒绝 | 拒绝 | 拒绝 | 拒绝 | 接受 |
| 拒绝年数 | 3 | 0 | 4 | 1 | 1 | 2 | 2 |

*来源：鹏元国际*

**表 24：七家评级公司信用评级在“AA”等级的三年期违约预测准确性二项式检验（2014 年-2017 年）**

| 年份 | 中诚信 | 中债 | 联合 | 中证鹏元 | 东方金诚 | 大公国际 | 上海新世纪 |
|---|---|---|---|---|---|---|---|
| 2014 | 拒绝 | 接受 | 拒绝 | 拒绝 | 接受 | 拒绝 | 拒绝 |
| 2015 | 拒绝 | 接受 | 拒绝 | 接受 | 拒绝 | 拒绝 | 接受 |
| 2016 | 拒绝 | 接受 | 拒绝 | 拒绝 | 拒绝 | 拒绝 | 拒绝 |
| 2017 | 拒绝 | 接受 | 拒绝 | 拒绝 | 拒绝 | 拒绝 | 拒绝 |
| 拒绝年数 | 4 | 0 | 4 | 3 | 3 | 4 | 3 |

*来源：鹏元国际*

**表 25：七家评级公司信用评级在“AA-”等级的三年期违约预测准确性二项式检验（2014 年-2017 年）**

| 年份 | 中诚信 | 中债 | 联合 | 中证鹏元 | 东方金诚 | 大公国际 | 上海新世纪 |
|---|---|---|---|---|---|---|---|
| 2014 | 拒绝 | 拒绝 | 拒绝 | 拒绝 | 接受 | 拒绝 | 拒绝 |
| 2015 | 接受 | 拒绝 | 接受 | 接受 | 拒绝 | 接受 | 接受 |
| 2016 | 接受 | 接受 | 接受 | 拒绝 | 接受 | 拒绝 | 拒绝 |
| 2017 | 接受 | 拒绝 | 接受 | 拒绝 | 接受 | 接受 | 接受 |
| 拒绝年数 | 1 | 3 | 1 | 3 | 1 | 2 | 2 |

*来源：鹏元国际*

**表 26：七家评级公司信用评级在“AA-”及以上的三年期违约预测准确性卡方检验（2014 年-2017 年）**

| 评级机构简称 | p 值 | | | | 验证结果 | | | | 拒绝年数 |
|---|---|---|---|---|---|---|---|---|---|
| | 2014 | 2015 | 2016 | 2017 | 2014 | 2015 | 2016 | 2017 | |
| 中诚信 | 0.53 | 0.40 | 0.63 | 0.83 | 接受 | 接受 | 接受 | 接受 | 0 |
| 中债 | 0.89 | 0.85 | 0.77 | 0.93 | 接受 | 接受 | 接受 | 接受 | 0 |
| 联合 | 0.37 | 0.72 | 0.00 | 0.09 | 接受 | 接受 | 拒绝 | 接受 | 1 |
| 中证鹏元 | 0.36 | 0.39 | 0.97 | 0.01 | 接受 | 接受 | 接受 | 拒绝 | 1 |
| 东方金诚 | 0.98 | 0.39 | 0.91 | 0.05 | 接受 | 接受 | 接受 | 接受 | 0 |
| 大公国际 | 0.04 | 0.68 | 0.05 | 0.09 | 拒绝 | 接受 | 接受 | 接受 | 1 |
| 上海新世纪 | 0.16 | 0.41 | 0.15 | 0.00 | 接受 | 接受 | 接受 | 拒绝 | 1 |

*来源：鹏元国际*

鵬元國際
PENGYUAN

## 免责声明

鹏元资信评估（香港）有限公司（“鹏元国际”、“鹏元”、“本公司”）按照既定的内部流程拟备不同的信用研究和相关评论（统称“研究”）。本公司保留在不事先通知的情况下，自行决定修改、更改、删除以及在其网站上发布任何资讯的权利。

研究适用于免责声明和限制。**研究和信用评级不是财务或投资建议，也不能被认爲是购买、出售或持有任何证券的建议，并且不能反映/针对任何证券的市场价值。我们认爲研究和信用评级的使用方应受过专业培训，有能力独立评估投资和商业决策。**

本研究完全根据本研究发表时作者可获得的公开数据和资讯作出。爲了本研究的目的，本公司会从我们认爲可靠且准确的来源获得足够有质量的事实性资讯。我们不会进行审计，也不会对研究中使用的任何资讯进行尽职调查或第三方校验。公司概不就研究中任何公开资讯的遗漏、错误或不一致性负责任。

**本公司不对以任何形式对其提供的任何资讯的准确性、及时性或完整性做出任何明示或暗示的保证。**在任何情况下，本公司、公司董事、股东、雇员、代表，均不对任何使用本公司发布的资讯所造成的的损害、开支、费用或损失承担任何责任。

本研究侧重于观察信用评级市场的发展趋势。本研究在向公衆发佈前，尚未提供给任何发行人。公司不会因其研究而获得酬金。

本公司保留其在公司网站、公司的社交媒体页面和授权第三方发佈研究的权利。未经本公司事先书面同意，不得以任何方式修改、複製、转载、传播或篡改本公司发布的任何内容。

本公司的研究并非给处在使用此研究可能构成违法的管辖区内的任何人传播或使用。如有疑问，请谘询相关的监管机构或专业顾问，以确保遵守适用的法律法规。

由于研究產生或与其相关的任何争议，本公司有权自行决定与争议解决相关的所有事宜，包括但不限于免责声明和政策的解读。