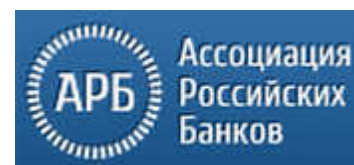


# Комитет по стандартам Базель II и управлению рисками



## Валидация

### Авторы:

Битюцкий Вячеслав,	Ernst&Young
Патратий Ольга,	ЮниКредит Банк
Перевицкая Вера,	Альфа-Банк
Писаренко Виталий,	ZEB
Чернышев Олег,	Ernst&Young

Москва, 2013

# Оглавление

Список сокращений.....	4
<b>I. Методические рекомендации по валидации внутренних моделей оценки вероятности дефолта клиента.....</b>	<b>5</b>
<b>1. Качественная валидация .....</b>	<b>7</b>
<b>2. Количественная валидация.....</b>	<b>8</b>
2.1. Тесты на дискриминационную способность .....	8
2.1.1. Оценка относительного ROPM.....	9
2.1.2. Оценка абсолютного ROPM .....	11
2.2. Калибровка модели .....	11
2.2.1. Тест на соответствие рассчитанного PD и наблюдаемого уровня дефолтов.....	12
2.2.3. Биномиальный тест.....	13
2.2.4. Тест хи-квадрат (Хосмер-Лемешев) .....	16
2.3. Репрезентативность/ Стабильность .....	16
2.4. Концентрация.....	17
2.5. Влияние процесса на результаты работы модели .....	17
<b>II. Методические рекомендации по валидации внутренних моделей оценки LGD и EAD .....</b>	<b>18</b>
<b>1. Первичная валидация LGD/EAD-моделей: Оценка модели пользователем .....</b>	<b>19</b>
1.1. Восприятие модели .....	19
1.2. Использование модели и степень интеграции в кредитный процесс.....	19
1.3. Требования к документации .....	20
<b>2. Продвинутая валидация LGD/EAD-моделей: Оценка результатов.....</b>	<b>20</b>
2.1. Оценка эффективности модели .....	20
2.2. Бэнчмаркинг.....	21
2.3. Построение тестовой выборки.....	23
<b>3. Комплексная валидация LGD/EAD-моделей: Качественное тестирование методики .....</b>	<b>24</b>
3.1. Анализ драйверов риска и особенностей расчета прогнозного значения LGD .....	24
3.1.1. Макроэкономические факторы.....	25
3.1.2. Параметры заемщика и характеристики кредитного продукта .....	25
3.1.3. Параметры обеспечения .....	25
3.1.4. Гарантии.....	26
3.1.5. Внутренние факторы банка .....	26
3.2. Анализ особенностей расчета реализовавшихся потерь .....	26
3.3. Анализ драйверов риска и особенностей расчета прогнозного значения EAD .....	26
<b>4. Комплексная валидация LGD/EAD-моделей: Качественное тестирование данных .....</b>	<b>27</b>
4.1. Анализ методов агрегации и построения исходной выборки .....	27
4.2. Проверка сопоставимости данных.....	27
4.3. Наличие в выборке данных, соответствующих экономическому спаду .....	28
<b>III. Методические рекомендации по валидации процесса присвоения рейтинга, использования риск-параметров и адекватности систем контроля.....</b>	<b>29</b>
<b>A. Валидация процесса присвоения рейтинга, использования риск-параметров и адекватности систем контроля.....</b>	<b>29</b>
<b>1. Процесс присвоения рейтингов.....</b>	<b>30</b>
1.1. Внутренние нормативные документы .....	30
1.2. Целостность .....	31
1.3. Однородность и уникальность.....	31
1.4. Обновление рейтинга и покрытие .....	31
1.5. Обучение .....	31
<b>2. Внутреннее использование .....</b>	<b>31</b>
2.1. Степень использования риск-параметров .....	31
2.2. Отчетность и мониторинг .....	32
<b>3. Система контроля.....</b>	<b>32</b>
3.1. Сегментация и ведение групп взаимосвязанных заемщиков .....	32
3.2. Качество данных, используемых для расчета рейтинга .....	32
<b>4. Одобрение системы пользователями.....</b>	<b>33</b>

<b>В.</b>	<b>Валидация процесса по снижению кредитного риска .....</b>	<b>33</b>
<b>1.</b>	<b>Юридическая однозначность .....</b>	<b>34</b>
<b>2.</b>	<b>Своевременная ликвидация .....</b>	<b>35</b>
<b>3.</b>	<b>Организационные требования .....</b>	<b>35</b>
	3.1. Общие требования .....	35
	3.1.1. Политики и процедуры .....	35
	3.1.2. Оценка и процесс мониторинга .....	35
	3.1.3. Управление данными .....	36
	3.1.4. Контроль .....	36
	3.1.5. Отчетность .....	36
	3.1.6. Организационная структура .....	36
	3.2. Специальные требования .....	36
	3.2.1. Фондированное кредитное обеспечение .....	36
	3.3. Кредитное обеспечение в виде гарантий/поручительств .....	37
<b>4.</b>	<b>Раскрытие информации.....</b>	<b>37</b>
<b>С.</b>	<b>Валидация информационных технологий, задействованных в рейтинговой системе .....</b>	<b>37</b>
<b>1.</b>	<b>Валидация ИТ, связанных с моделью .....</b>	<b>38</b>
	1.1. Среда разработки .....	38
	1.2. Валидация ИТ платформы .....	39
	1.3. Промышленная среда .....	39
	<b>Приложения.....</b>	<b>41</b>
<b>1.</b>	<b>Пороговые значения для интерпретации результатов количественных тестов .....</b>	<b>41</b>
<b>2.</b>	<b>Показатели качественной валидации .....</b>	<b>45</b>
	2.1. AUROC – площадь под ROC кривой .....	45
	2.2. Мера AR (Accuracy Ratio) .....	47
	2.3. Оценка стандартной ошибки для индекса Джини .....	48
	2.4. Критерий Колмогорова-Смирнова .....	50
	2.5. Мера Information value (IV).....	50
<b>3.</b>	<b>Инструменты оценки стабильности моделей LGD/EAD .....</b>	<b>52</b>
	3.1. Индекс стабильности системы.....	52
	3.1.1. Критерий согласия Колмогорова-Смирнова .....	52
	3.2. Инструменты оценки дискриминационной способности моделей LGD .....	53
	3.2.1. Loss Capture Ratio – Коэффициент покрытия убытков .....	53
	3.2.2. Cumulative LGD Accuracy Ratio (CLAR) – кумулятивный коэффициент точности LGD.....	54
	3.2.3. Коэффициент ранговой корреляции Спирмена.....	54
	3.3. Инструменты оценки качества калибровки моделей LGD/EAD.....	55
	3.3.1. Loss Shortfall LGD.....	55
	3.3.2. Стандартное абсолютное отклонение LGD.....	56
	3.3.3. Калибровка LGD/EAD при помощи критерия Стьюдента.....	56
<b>4.</b>	<b>Пороговые значения для интерпретации результатов количественных тестов .....</b>	<b>57</b>

## Список сокращений

- AR (accuracy ratio) – коэффициент точности
- AUROC (area under receiver operator characteristic) – площадь под ROC кривой
- CAP (cumulative accuracy profile) – совокупный профиль точности
- CRM (credit risk mitigation) – процесс снижения кредитного риска
- CT (central tendency) – центральная тенденция
- EAD (exposure at default) – величина кредитного требования, подверженная риску дефолта.
- IV (information value) – значимость информации
- LGD (loss given default) – уровень потерь при дефолте.
- PD (probability of default) – вероятность дефолта
- PIT (point in time) – калибровка «на момент времени»
- PSI (population stability index) – индекс стабильности популяции
- ROC curve (receiver operator characteristic) – кривая соотношения определения «хороших» и «плохих» наблюдений
- ROPM (rank ordering performance measure) – мера качества ранжирования
- TTC (through the cycle) – калибровка «по циклу»

## I. Методические рекомендации по валидации внутренних моделей оценки вероятности дефолта клиента

Цель валидации заключается в подтверждении соответствия рейтинговой системы нормативным требованиям и адекватности вычисляемых риск – параметров.

Документ содержит информацию о тестах, позволяющих оценить качество внутренних моделей оценки вероятности дефолта клиентов высокодефолтных портфелей и является приложением к Методологическим рекомендациям по расчету кредитного риска на основе внутренних рейтингов.

Базельским Комитетом были сформулированы **принципы проведения валидации**<sup>1</sup>:

- В первую очередь банк является ответственным за проведение валидации;
- Валидация, по своей сути, является оценкой предсказательной способности риск - параметров, рассчитанных банком, а также их применения в кредитном процессе;
- Валидация – итеративный процесс;
- Не существует единого подхода по проведению валидации;
- Валидация должна включать как количественную, так и качественную оценку;
- Следует проводить независимую проверку процесса и результатов валидации.

В документе представлены различные подходы по валидации PD моделей, рекомендации по выполнению тестов, а также необходимая справка из теории статистики. Формат отчета о валидации, последовательность выполнения тестов, их объем, тесты для возможных дополнительных исследований определяются банком самостоятельно.

Валидация PD моделей подразумевает как количественную, так и качественную оценку модели.

**Качественная валидация** должна включать оценку следующих аспектов:

- Документация модели
- Методология разработки модели
- Определение дефолта
- Глубина и качество данных, использованных для разработки
- Подход к калибровке
- Применяемая рейтинговая шкала
- Применяемые корректировки рассчитанного рейтинга

**Количественная валидация** PD моделей подразумевает оценку следующих аспектов:

- Дискриминационная способность модели
- Калибровка модели
- Репрезентативность
- Концентрация
- Стабильность
- Влияние процесса на результаты работы модели

---

<sup>1</sup> Принципы проведения валидации, а также возможные подходы изложены в документе ‘Studies on the validation of internal rating systems, revised version May 2’ подготовленном Базельским комитетом по банковскому надзору.

Документ содержит варианты тестов для трех типов валидации, различающихся по целям и глубине анализа. Таким образом, для каждой категории тестов указано, в каком случае необходимо их выполнение. Различают **3 типа валидации**:

- *Первоначальная* - валидация, проводимая при разработке/до ввода в действие модели, а также в случае существенных<sup>2</sup> изменений модели. Данный тип валидации включает в себя тесты по структуре модели и тесты на эффективность модели, оценивает ее соответствие, по меньшей мере, минимальным требованиям.

- *Мониторинг* - с частой периодичностью проводимая валидация (например, ежемесячная, ежеквартальная), включающая в себя проведение тестов высокого уровня и направленная на выявление ранних признаков ухудшения эффективности модели.

- *Последующая* - ежегодная углубленная валидация, включающая проведение всех групп тестов для подтверждения адекватного функционирования модели в случае отсутствия существенных изменений в модели. Разработчики модели вправе запросить внеплановую последующую валидацию в случае выявления признаков значительного ухудшения модели в рамках процесса мониторинга.

Таблица 1

### Типы валидации и группы тестов

Группы тестов	Первоначальная валидация	Мониторинг	Последующая валидация
<i>Качественная валидация</i>			
<ul style="list-style-type: none"> <li>• Документация</li> <li>• Методология разработки</li> <li>• Данные для разработки</li> <li>• Подход к калибровке</li> <li>• Рейтинговая шкала</li> </ul>	При разработке/ При существенных изменениях модели		
<ul style="list-style-type: none"> <li>• Определение дефолта для калибровки модели</li> <li>• Корректировки рассчитанного рейтинга</li> </ul>	При разработке/ При существенных изменениях модели		Ежегодно
<i>Количественная валидация</i>			
<ul style="list-style-type: none"> <li>• Дискриминационная способность модели в целом/ модулей/ сегментов/ факторов</li> <li>• Калибровка</li> <li>• Концентрация / дисперсия</li> <li>• Репрезентативность</li> <li>• Стабильность</li> </ul>	При разработке/ При существенных изменениях модели	Ежемесячно/ Ежеквартально	Ежегодно
<ul style="list-style-type: none"> <li>• Влияние процесса на модель</li> </ul>		Ежемесячно/ Ежеквартально	Ежегодно

<sup>2</sup> Критерии существенности должны быть определены в нормативных документах.

## 1. Качественная валидация

Тесты по структуре и методологии модели направлены на оценку качества и адекватности выбора подходов построения модели, их сравнения с подходами, использованными при построении аналогичных моделей, а также в случае, если произошли изменения в составе портфеля клиентов /экономической среды, которые могут привести к необходимости пересмотра выбранных подходов.

В рамках качественной валидации модели, целесообразно детально проанализировать следующие аспекты:

Тема	Рассматриваемые вопросы
<i>Общие принципы</i>	
<p>Качество документации по разработке модели. Возможно ли по имеющимся документам в точности воспроизвести рассматриваемую модель.</p>	<ul style="list-style-type: none"> <li>• Представлены ли все этапы разработки модели.</li> <li>• Логичность документации.</li> <li>• Возможность воспроизведения модели по документации.</li> <li>• Корректность количественных показателей в отчете.</li> <li>• Наличие в документации обоснования сделанных в ходе разработке допущений.</li> <li>• Наличие в документации информации об известных недостатках модели и случаях, при которых ее использование неэффективно.</li> </ul>
<p>Позволяет ли выбранный подход к сегментации корректно разделять клиентов по профилю риска.</p>	<ul style="list-style-type: none"> <li>• Достаточность обоснования необходимости применения сегментации.</li> <li>• Согласован ли выбранный подход к сегментации со всеми заинтересованными лицами.</li> <li>• Проанализирована ли ситуация появления новых сегментов.</li> <li>• Отсутствие возможности манипулирования сегментом клиента.</li> </ul>
<p>Учитывалось ли мнение экспертов при разработке модели.</p>	<ul style="list-style-type: none"> <li>• Учитывалось ли мнение экспертов при принятии существенных решений по структуре модели (определение сегментов, выбор факторов и т.п.)</li> <li>• Наличие протоколов по результатам совместной с экспертами работы.</li> </ul>
<i>Методология разработки модели</i>	
<p>Соответствует ли определение дефолта, используемое при разработке модели нормативным требованиям.</p>	<ul style="list-style-type: none"> <li>• Соответствие определения дефолта, применяемое для разработки модели нормативным требованиям.</li> <li>• Используемый критерий материальности.</li> </ul>
<p>Соответствует ли методология разработки структуре и размеру портфеля, адекватен ли подход к сегментации, обоснованы ли все допущения и предположения.</p>	<ul style="list-style-type: none"> <li>• Соответствие выбранного подхода разработки модели общепринятым стандартам для рассматриваемого портфеля.</li> <li>• Все ли возможные модули для оценки кредитоспособности заемщика были рассмотрены/разработаны.</li> <li>• Были ли рассмотрены альтернативные варианты комбинации модулей/ калибровки.</li> </ul>
<p>Все ли возможные источники данных использовались при разработке, соответствует ли глубина данных нормативным требованиям.</p>	<ul style="list-style-type: none"> <li>• Все ли доступные источники данных были использованы при разработке.</li> <li>• Подтверждено ли качество использованных при разработке данных ответственными лицами.</li> <li>• Наличие документации с описанием алгоритмов выгрузки и обработки данных для разработки.</li> <li>• Соответствие глубины исторических данных нормативным требованиям.</li> </ul>
<p>Являлись ли критерии выбора факторов модели исчерпывающими.</p>	<ul style="list-style-type: none"> <li>• Является ли первоначальный список факторов исчерпывающим.</li> <li>• Соответствует ли алгоритм выбора факторов из первоначального списка общепринятым практикам (например: одномерный анализ, корреляционный анализ, многомерный анализ).</li> </ul>

Тема	Рассматриваемые вопросы
	<ul style="list-style-type: none"> <li>• Качество факторов, вошедших в финальную модель (например, отсутствие высоких корреляций, статистическая значимость факторов).</li> </ul>
<i>Подход к калибровке</i>	
Выборка для калибровки модели и используемое определение дефолта. Выбранный подход к калибровке.	<ul style="list-style-type: none"> <li>• Отличие выборки для калибровки от выборки для построения модели.</li> <li>• Источники и глубина данных, использованных для калибровки.</li> <li>• Соответствие определения дефолта, использованного для калибровки нормативным требованиям.</li> <li>• Соответствие разработанной калибровки модели выбранному подходу.</li> </ul>
Рейтинговая шкала, ее адекватность и соответствие требованиям регулятора.	<ul style="list-style-type: none"> <li>• Подход по выбору рейтинговой шкалы, его соответствие требованиям регулятора.</li> <li>• Отсутствие избыточных концентраций по интервалам рейтинговой шкалы.</li> <li>• Адекватность распределения клиентов по интервалам рейтинговой шкалы.</li> </ul>
<i>Корректировки рейтинга</i>	
Учет влияния групп взаимосвязанных заемщиков.	<ul style="list-style-type: none"> <li>• Учет внутригрупповых связей в модели.</li> <li>• Влияние на качество модели подобных корректировок.</li> </ul>
Учет влияния предупредительных сигналов.	<ul style="list-style-type: none"> <li>• Наличие модуля предупредительных сигналов.</li> <li>• Оценка влияния такого модуля на общее качество работы модели.</li> </ul>
Существует ли возможность изменить рассчитанный рейтинг вручную.	<ul style="list-style-type: none"> <li>• Случаи возможного ручного изменения финального рассчитанного рейтинга.</li> <li>• Обоснованность причин изменения рассчитанного рейтинга.</li> <li>• Адекватность установленных порогов по ручной корректировке рейтинга.</li> <li>• Влияние ручных корректировок на качество работы модели.</li> </ul>
<i>Добавка консерватизма в модели</i>	
Наличие поправок, добавляющих модели консерватизма.	<ul style="list-style-type: none"> <li>• Были ли при разработке модели включены поправки, добавляющие ей консерватизма.</li> <li>• Перечень недостатков, покрываемый этой поправкой.</li> <li>• Наличие обоснования и алгоритма расчета поправки, добавляющей модели консерватизма.</li> </ul>
<i>Соответствие качества модели требованиям в течение периода использования (use test)</i>	
Обзор модели с момента внедрения на соответствие нормативным требованиям, предъявляемым в период использования (use test).	<ul style="list-style-type: none"> <li>• Вносились ли в модель изменения с момента внедрения, их существенность.</li> <li>• Изменение дискриминационной способности модели с момента внедрения во времени.</li> </ul>

## 2. Количественная валидация

В данном документе результаты количественной валидации предполагается оценивать по «принципу светофора», как наиболее широко используемому в банковском регулировании.

### 2.1. Тесты на дискриминационную способность

Тесты на оценку качества ранжирования клиентов моделью необходимо проводить по следующим направлениям:

- **общая оценка эффективности ранжирования модели (ROPM)** – необходимо для понимания качества модели в целом;
- **оценка эффективности ранжирования модели в разрезе сегментов портфеля** – позволяет оценить, на сколько эффективно модель работает для каждого из сегментов модели;



- эффективность ранжирования модулей/компонентов/отдельных факторов модели – необходимо проводить такую оценку, даже если в целом модель работает на приемлемом уровне. Подобный анализ позволяет выявить как возможные ошибки в модели, так и источники снижения качества работы модели.

Тесты следует применять как на относительном уровне (т.е. должна проводиться оценка изменения эффективности ранжирования между валидационной выборкой и выборкой, использованной при разработке), так и на абсолютном уровне (т.е. в сравнении абсолютного уровня эффективности ранжирования с установленными регулятором значениями).

Необходимо отметить, что методологии, используемые для оценки всех трех компонентов, имеют между собой много общего. В частности, тесты для общей оценки эффективности ранжирования модели также используются для оценки эффективности ранжирования отдельных факторов или модулей/компонентов модели.

Качество модели PD существенно зависит от ее способности отделять «хороших» клиентов (с низким PD) от плохих (с высоким PD) путем присвоения скорингового балла каждому клиенту.

Для целей проведения валидации необходимо использовать те же меры оценки эффективности ранжирования модели, что и при ее разработке. Например, можно использовать Accuracy Ratio для портфелей с высоким уровнем дефолта и Sommers'D – для портфелей с низкой вероятностью дефолта. В Приложении 2 приведены детальные описания возможных методов оценки качества ранжирования наблюдений/ клиентов моделью: AUROC, AR (Accuracy Ratio, Gini), критерий Колмогорова-Смирнова, Information value, тем не менее необходимо принимать во внимание соответствующие ограничения по применению этих методов в т.ч. по количеству наблюдений. Также, для оценки ROPM часто используются такие методы как Sommers'D, Kendall's Tau и прочие.

### 2.1.1. Оценка относительного ROPM

Как правило, качество ранжирования клиентов моделью снижается с момента разработки к моменту валидации. Таким образом, масштаб снижения дискриминационной способности следует оценивать на предмет приемлемости по «**принципу светофора**»:

- «**Зеленый**» результат – незначительное/отсутствие снижения ROPM, разница находится в приемлемых границах.
- «**Желтый**» результат – снижение ROPM достаточно существенное для обозначения необходимости проведения более глубокого анализа, но разница все еще находится в приемлемых границах.
- «**Красный**» результат – снижение ROPM значительно, выпадает из приемлемых границ, требуется глубокий анализ причин существенных изменений.

Оценка приемлемости результата отражает непосредственно степень изменения ROPM и не зависит от размера выборки, в то же время размер выборки является существенным фактором, позволяющим определить, насколько достоверна полученная оценка. В этом случае, каждую оценку приемлемости изменения ROPM, необходимо сопровождать оценкой достоверности полученного результата. Уровень достоверности результата может принимать одно из четырех значений:

- **Высокий уровень** – полученный уровень приемлемости очень близок к действительному, что доказывается статистически.
- **Средний уровень** – полученный уровень приемлемости, скорее всего, близок к действительному.
- **Низкий уровень** – близость полученного уровня приемлемости к действительному статистически подтверждается очень слабо
- **Неопределенный** – вывод о близости полученного уровня приемлемости к действительному статистически подтвердить не представляется возможным.

Таким образом, имеющаяся оценка приемлемости результата и ее достоверности, позволяет сделать вывод об изменении дискриминационной способности модели с момента разработки до момента валидации.

Если присутствует снижение дискриминационной способности модели в целом – это может являться следствием следующих причин:

- Дискриминационная способность всех факторов практически не изменилась, но подход по включению их в модель более не приемлем (т.е. веса факторов не оптимальны).
- Дискриминационная способность всех факторов изменилась более-менее однородно.
- Дискриминационная способность одного или нескольких факторов резко изменилась.

Существует несколько различных путей повышения качества модели, связанных с первыми двумя причинами. Однако, в случае значительного снижения ROPM и низкого качества модели, наилучшим методом является ее переработка. При наличии проблем с одним/ несколькими факторами, возможно доработать/ изменить только их. Таким образом, оценку ROPM на общем уровне модели всегда следует сопровождать оценкой на уровне модулей/ сегментов / отдельных факторов.

Для того чтобы оценить относительное изменение дискриминационной способности, необходимо рассчитать следующие параметры:

- $ROPM_D$  – значение ROPM для модели/ компонента модели на выборке разработки;
- $ROPM_V$  - значение ROPM для модели/ компонента модели на выборке валидации;
- $\sigma_D^{ROPM}$  – стандартная ошибка, соответствующая  $ROPM_D$ ,
- $\sigma_V^{ROPM}$  - стандартная ошибка, соответствующая  $ROPM_V$ ,
- $CutOff_{yellow}$  – граничное изменение ROPM, превышение которого соответствует «желтому» результату прохождения теста,
- $CutOff_{red}$  – граничное изменение ROPM, превышение которого соответствует «красному» результату прохождения теста,
- $\alpha_H, \alpha_M, \alpha_L$  – высокий, средний и низкий уровни значимости для оценки достоверности полученных результатов.

Таким образом, изменение ROPM оценивается по следующей формуле:

$$ROPM_V - ROPM_D + CutOff_{yellow} > 0 \rightarrow \text{результат Зеленый}$$

$$ROPM_V - ROPM_D + CutOff_{red} > 0 \rightarrow \text{результат Желтый}$$

$$ROPM_V - ROPM_D + CutOff_{red} < 0 \rightarrow \text{результат Красный}$$

Для оценки достоверности результатов, предполагается, что ROPM независимы и имеют нормальное распределение<sup>3</sup>, и рассчитываются две статистики:

$$T_{yellow} = \frac{ROPM_V - ROPM_D + CutOff_{yellow}}{\sqrt{(\sigma_D^{ROPM})^2 + (\sigma_V^{ROPM})^2}}$$

$$T_{red} = \frac{ROPM_V - ROPM_D + CutOff_{red}}{\sqrt{(\sigma_D^{ROPM})^2 + (\sigma_V^{ROPM})^2}}$$

---

<sup>3</sup> Это утверждение обосновано для оценок AR и AUROC, но примерно верно для критерия Колмогорова-Смирнова и, в частности, для IV (Information Value).

Задавая один из уровней значимости, можно определить, на сколько полученный результат является достоверным:

- $T_{yellow} < \Phi^{-1}(\alpha)$  – для заданного уровня значимости результат не «зеленый»,
- $T_{yellow} > \Phi^{-1}(1 - \alpha)$  – для заданного уровня значимости результат не является ни «желтым», ни «красным»,
- $T_{red} > \Phi^{-1}(1 - \alpha)$  – для заданного уровня значимости результат не является «красным»,
- $T_{red} < \Phi^{-1}(\alpha)$  – для заданного уровня значимости, результат не является ни «зеленым», ни «желтым».

В случае если перечисленные условия не выполняются ни для одного из заданных уровней значимости, достоверность результата неопределенна.

Таким образом, полученные результаты позволяют сделать вывод об относительном уровне снижения ROPM и его достоверности.

Иногда в рамках валидации выявляются случаи существенного улучшения дискриминационной способности отдельных факторов / модулей. Источником улучшений может являться как изменение структуры портфеля, так и изменение его качества. В любом случае, банк самостоятельно определяет необходимость проведения дополнительных исследований в части репрезентативности портфеля, оценки влияния изменений на модель в целом и т.д.

### 2.1.2. Оценка абсолютного ROPM

Результаты оценки ROPM можно подразделить на три категории по «принципу светофора»:

- «**Зеленый**» – ROPM находится в приемлемых границах.
- «**Желтый**» – ROPM недостаточно высокий, но все еще находится в допустимых границах, возможно требуются дополнительные исследования.
- «**Красный**» – ROPM находится на неудовлетворительном уровне, качество модели/ компонента модели сложно подтвердить, обязательны дополнительные исследования.

Аналогично случаю с относительным значением ROPM, кроме непосредственно оценки ROPM, важна и достоверность полученного результата. Достоверность может быть высокая, средняя, низкая и может быть не определена. Достоверность результата может быть проверена теми же методами, что использовались ранее (см. гл. 2.1.1.).

В Приложении 2 представлены различные подходы по оценке ROPM.

## 2.2. Калибровка модели

Не все рейтинговые модели возвращают PD, которые совпадут с фактическим уровнем дефолта в будущем году. Связь вероятности дефолта и наблюдаемым фактическим уровнем дефолта зависит от выбранного подхода к калибровке. Различают три подхода к калибровке:

- Подход «**на момент времени**» (point in time - PIT) - ожидается, что PD портфеля и наблюдаемый уровень дефолта будут близкими в каждом рассматриваемом году.
- Подход «**по циклу**» (through the cycle - TTC) - ожидается, что средний PD портфеля будет близким к среднему за цикл уровню дефолта (central tendency, CT); т.о. в периоды экономического роста/ спада PD портфеля будет отличаться от наблюдаемого уровня дефолта.
- **Стандартная (смешанная)** калибровка – PD портфеля лишь частично следует циклу. Так, в период экономического спада предсказываемый PD портфеля будет ниже наблюдаемого уровня дефолта, а в период роста – выше.

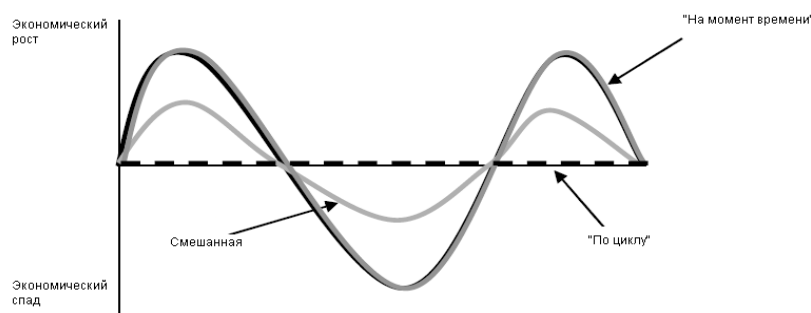


Рис 4. Различные подходы к калибровке

Ниже будут рассмотрены некоторые тесты, которые могут использоваться для проверки качества калибровки модели.

### 2.2.1. Тест на соответствие рассчитанного PD и наблюдаемого уровня дефолтов

Первым шагом при проверке качества калибровки необходимо сравнить фактический уровень дефолта и рассчитанную вероятность дефолта по портфелю. Сравнение нужно проводить с учетом выбранного подхода к калибровке. Так, в случае подхода «на момент времени» можно ожидать, что разница между уровнем дефолта и PD не должна быть существенной, то для подхода «по циклу» разница может оказаться значительной в силу, например, высокой волатильности экономики.

Допустимые границы отклонений рассчитанного PD и уровня дефолтов, а также методика сравнения в случае смешанного подхода по калибровке или ТТС, должны определяться банком самостоятельно.

Тем не менее, стоит отметить, что неудовлетворительные результаты теста не всегда свидетельствуют о низком качестве калибровки. В такой ситуации необходимо провести дополнительный углубленный анализ для выяснения причин отрицательного результата теста.

Тест на соответствие рассчитанного PD и фактического уровня дефолтов проводят в следующих разрезах:

- На уровне портфеля в целом
- На уровне сегментов

### 2.2.2. Тест на точность калибровки модели

Цель теста – проверить калибровку модели: насколько фактические дефолты соответствуют прогнозным дефолтам в целом по всей выборке. Суть теста состоит в проверке попадания наблюдаемого уровня дефолтов в доверительный интервал вокруг оценки вероятности дефолта.

Строятся три интервала:  $CI_{95\%}$  (уровень значимости 0.05),  $CI_{99\%}$  (уровень значимости 0.01) и минимальный интервал. Алгоритм их построения следующий.

Вероятность того, что в  $n$  наблюдениях произойдет  $q$  дефолтов, в соответствии с биномиальным законом распределения равна:

$$P(Y = q, n) = \binom{n}{q} PD^q (1 - PD)^{n-q},$$

где  $Y$  – наблюдаемое количество дефолтов,  $PD$  – вероятность дефолта, полученная с помощью модели,  $\binom{n}{q} = \frac{n!}{(n-q)!q!}$

Тогда вероятность того, что наблюдаемое количество дефолтов  $Y$  не превысит  $u$ , равна

$$P(Y \leq y, n) = \sum_{q=0}^{\lfloor y \rfloor} \binom{n}{q} PD^q (1 - PD)^{n-q},$$

где  $\lfloor y \rfloor$  – наибольшее целое число, не превосходящее  $y$ .

В соответствии с данной формулой выбираются значения  $y_l^{0.95}, y_h^{0.95}$  – нижняя и верхняя границы 95%-ного доверительного интервала таким образом, что

$$\inf\{y_l^{0.95} | P(Y \leq y_l^{0.95}, n) \geq 0.025\} \text{ и } \inf\{y_h^{0.95} | P(Y \leq y_h^{0.95}, n) \geq 0.975\}.$$

Аналогично для 99%-ного интервала

$$\inf\{y_l^{0.99} | P(Y \leq y_l^{0.99}, n) \geq 0.005\} \text{ и } \inf\{y_h^{0.99} | P(Y \leq y_h^{0.99}, n) \geq 0.995\}.$$

Доверительные интервалы для вероятности дефолтов получаются делением границ доверительных интервалов для количества дефолтов на общее количество наблюдений.

Минимальный интервал  $CI_{min}$  строится для того, чтобы в случае очень большого количества имеющихся наблюдений заложить минимальный допустимый уровень отклонения  $DR$  от  $PD$  в случае, когда такое отклонение не оказывает существенного влияния на уровень рисков банка вследствие малого значения отклонения и малых значений уровней дефолтов (например в случае портфеля ипотечных или авто-кредитов, где размер портфеля может быть существенным в то время как уровень потерь незначителен).

*Интерпретация результатов.* Оценивается факт попадания фактического уровня дефолтов ( $DR$ ) в построенные интервалы.

Возможны три варианта соотношений трех интервалов:

- 1)  $CI_{min} \not\subseteq CI_{95\%}$
- 2)  $CI_{95\%} \subseteq CI_{min} \subseteq CI_{99\%}$
- 3)  $CI_{min} \not\supseteq CI_{99\%}$ .

### Интерпретация теста

Цвет светофора	Вариант		
	1	2	3
	$DR \notin CI_{99\%}$	$DR \notin CI_{99\%}$	$DR \notin CI_{min}$
	$DR \notin CI_{95\%}$ и $DR \in CI_{99\%}$	$DR \notin CI_{min}$ и $DR \in CI_{99\%}$	
	$DR \in CI_{95\%}$	$DR \in CI_{min}$	$DR \in CI_{min}$

Также проводится тест на правильность наклона калибровочной кривой. Для этого наблюдения упорядочиваются по величине скорингового балла и затем разбиваются на группы, соответствующие диапазонам скорингового балла. Затем для каждой группы проводится тест, аналогичный описанному выше, на попадание наблюдаемого уровня дефолтов в доверительный интервал вокруг оценки вероятности дефолта.

### 2.2.3. Биномиальный тест

Следующим этапом проверки качества калибровки является проверка соответствия рассчитанного значения  $PD$  и фактического уровня дефолтов для каждого интервала рейтинговой шкалы (т.е. насколько точна калибровочная кривая).

Сравнительный анализ проводится по следующему алгоритму:

- Для каждого среднего  $PD$  по интервалам рейтинговой шкалы рассчитываются доверительные интервалы. Для этого необходимо:

- Количество наблюдений в рассматриваемом интервале шкалы;
- Количество дефолтных наблюдений в рассматриваемом интервале шкалы;
- Уровень значимости для ширины доверительного интервала;
- Уровень допустимого отличия для придания некоторой значимой ширины доверительному интервалу в случае большого объема имеющихся данных (иначе ширина доверительного интервала будет близка к нулю);

- Оценивается, попадает ли имеющийся уровень дефолта в рассчитанный доверительный интервал;

- В зависимости от выбранного уровня значимости, возможно некоторое количество отклонений, даже если калибровочная кривая построена корректно. Ожидаемое количество отклонений рассчитывается и отнимается от имеющегося количества отклонений;

- Результат прохождений теста определяется по предварительно заданному количеству отклонений;

Ниже приведены необходимые формулы для выполнения этого теста.

$$PD = \frac{\sum_i N_i \cdot PD_i}{N}, DR_i = \frac{D_i}{N_i},$$

где  $N = \sum_i N_i$  – общее количество наблюдений

$N_i$  – количество наблюдений в  $i$ -ом интервале рейтинговой шкалы

$D_i$  – количество дефолтов, соответствующее  $i$ -ому интервалу рейтинговой шкалы

Установив уровень значимости  $\alpha$  и количество допустимых отклонений  $Disc$ , можно посчитать минимальные и максимальные значения уровней дефолта для каждого интервала рейтинговой шкалы по следующей формуле:

$$DR_i^{min,\alpha} = \frac{Binom^{-1,min} \left( N_i, DR_i^{low}, \frac{\alpha}{2} \right)}{N_i} \approx \frac{Poisson^{-1,min} \left( N_i \cdot DR_i^{low}, \frac{\alpha}{2} \right)}{N_i}$$

$$\approx \frac{int \left( N_i \cdot \left( DR_i^{low} + \sqrt{\frac{DR_i^{low} \cdot (1 - DR_i^{low})}{N_i}} \cdot \Phi^{-1} \left( \frac{\alpha}{2} \right) \right) \right)}{N_i}$$

$$DR_i^{max,\alpha} = \frac{Binom^{-1,max} \left( N_i, DR_i^{high}, 1 - \frac{\alpha}{2} \right)}{N_i} \approx \frac{Poisson^{-1,max} \left( N_i \cdot DR_i^{high}, 1 - \frac{\alpha}{2} \right)}{N_i}$$

$$\approx \frac{int \left( N_i \cdot \left( DR_i^{high} + \sqrt{\frac{DR_i^{high} \cdot (1 - DR_i^{high})}{N_i}} \cdot \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) \right) \right)}{N_i}$$

где

$Binom^{-1}$  – обратная функция кумулятивного биномиального распределения. Так как биномиальное распределение является дискретным, то существует два целочисленных варианта для обратного распределения:  $Binom^{-1,max}$  является наибольшим,  $Binom^{-1,min}$  – наименьшим.

$Poisson^{-1}$  – обратная функция кумулятивного распределения Пуассона. Так как распределение Пуассона является дискретным, то существует два целочисленных варианта для обратного распределения:  $Poisson^{-1,max}$  является наибольшим,  $Poisson^{-1,min}$  – наименьшим.

$\Phi^{-1}$  – обратное кумулятивное нормальное распределение. Нужно обратить внимание, что результаты округляются до целого числа, для того, чтобы  $DR_i^{max,\alpha}$  и  $DR_i^{min,\alpha}$  соответствовали фактическому уровню дефолтов. Округление выполняется для расширения допустимого диапазона.

$DR_i^{high} = \min(1, (1 + Disc) \cdot PD_i)$  – верхняя граница допустимого интервала для фактического уровня дефолта.

$DR_i^{low} = (1 - Disc) \cdot PD_i$  – нижняя граница допустимого интервала для фактического уровня дефолта.

Необходимо отметить, что биномиальное распределение может быть аппроксимировано распределением Пуассона (если  $N_i$  велико и  $PD_i$  мало) или нормальным распределением (если  $N_i \cdot PD_i > 10^{15}$  и  $(1 - PD_i) \cdot N_i > 10$ ).

Если для рассматриваемого интервала рейтинговой шкалы, фактический уровень дефолта выпадает из интервала, определяемого  $DR_i^{max,\alpha}$  и  $DR_i^{min,\alpha}$ , то это будет являться отклонением. Общее количество отклонений по всем интервалам рейтинговой шкалы назовем  $Dev$ , а количество интервалов –  $J$ . Тогда, количество отклонений, по которому оценивается качество калибровочной кривой, рассчитывается как:

$$Dev^{excess} = Dev - \alpha \cdot J$$

В зависимости от количества отклонений, результаты теста могут быть следующими:

$$\frac{Dev^{excess}}{J} < Cut_{Yellow}^{rel} \quad \text{результат Зеленый}$$

$$Cut_{Yellow}^{rel} \leq \frac{Dev^{excess}}{J} < Cut_{Red}^{rel} \quad \text{результат Желтый}$$

$$\frac{Dev^{excess}}{J} \geq Cut_{Red}^{rel} \quad \text{результат Красный}$$

где  $Cut_{Yellow}^{rel}$  и  $Cut_{Red}^{rel}$  - пороговые значения для долей возможных отклонений, превышение которых приводит к «желтому» или «красному» результату соответственно.

Для каждого из  $i$ -го интервала рейтинговой шкалы вычисляется отн. ошибка  $err_i^{rel}$ :

$$err_i^{rel} = \frac{DR_i - PD_i}{PD_i}$$

И соответствующая автокорреляция:

$$\rho^{auto} = \rho(err_i^{rel}, err_{i-1}^{rel}).$$

Неудовлетворительные результаты выполнения теста могут быть получены в случаях:

- Существенного отличия PD от фактического уровня дефолтов в силу, например, подхода к калибровке.
- Калибровочная кривая не соответствует фактической ситуации.

Таким образом, если для калибровки был выбран подход «по циклу», и биномиальный тест показал отрицательный результат, сделать вывод о качестве калибровочной кривой может быть затруднительно. В данной ситуации целесообразно выполнить корректировку PD, учитывая заложную цикличность.

Биномиальный тест необходимо выполнять:

- на уровне портфеля в целом.
- на уровне сегментов.
- на уровне модулей (при наличии PD на уровне модуля).

При необходимости проведения дополнительных исследований могут быть выполнены также такие тесты как Хи-квадрат, подход Мертона и другие. Перечень дополнительных тестов может быть установлен как банком самостоятельно, например, по отношению к определенным портфелям, так и сотрудником, проводящим валидацию, применительно к исследуемой модели.

#### 2.2.4. Тест хи-квадрат (Хосмер-Лемешев)

Тест Хи-квадрат позволяет проверять нулевую гипотезу о совпадении распределения событий с некоторым заданным распределением. События предполагаются независимыми и одинаково распределенными, результаты каждого из событий должны быть взаимоисключающими. Хи-квадрат вычисляется следующим образом:

$$T_k = \sum_{i=0}^k \frac{(n_i p_i - \theta_i)^2}{n_i p_i (1 - p_i)}$$

где  $p_0, \dots, p_k$  – оцененные вероятности дефолта в  $K+1$  рейтинговом интервале,

$n_i$  – количество заемщиков в рейтинговом интервале  $i$

$\theta_i$  – количество заемщиков в дефолте в рейтинговом интервале  $i$ .

В соответствии с центральной предельной теоремой при  $n_i \rightarrow 0$  для всех  $i$ ,  $T_k$  стремится к распределению  $\chi_{k+1}^2$ , если все  $p_i$  являются действительными вероятностями дефолта.

Значение p-value теста Хи-квадрат может служить как мера точности оценки вероятности дефолта: чем ближе p-value к нулю, тем хуже оценка.

В случаях, когда ожидаемое значение в знаменателе оказывается малым (что означает либо малую вероятность, либо малое количество наблюдений), нормальная аппроксимация полиномиального распределения может не работать, тогда, возможно, стоит использовать G-тест, статистика, основанная на отношении правдоподобия.

Банкам предлагается самостоятельно устанавливать пороговые значения теста Хосмера-Лемешева, при этом рекомендуется учесть используемый подход к калибровке, а также волатильность уровня дефолта по портфелю.

### 2.3. Репрезентативность/ Стабильность

Тесты на репрезентативность направлены на подтверждение соответствия текущего состава портфеля выборкам, использованным для разработки и валидации модели. Существенные изменения в составе портфеля могут привести к значительному снижению качества работы модели и, как следствие, необходимости ее переработки. Для оценки репрезентативности используются характеристики клиентов, входящих в портфель.

В качестве меры изменения портфеля используется Population Stability Index (PSI), формула для расчета которого была приведена в Приложении 2 (см. гл. Information Value). Предлагаемые допустимые границы изменений приведены в Приложении 1.

Тесты на репрезентативность необходимо проводить в следующих разрезах:

- Соответствие выборки для разработки текущему портфелю.
- Соответствие выборки для калибровки текущему портфелю.
- Соответствие выборки для валидации текущему портфелю.

Тесты на стабильность выполняются для выявления изменений в характеристиках портфеля. Таким образом, для выполнения теста необходимо проводить сравнение между выборками для разработки и валидации в следующих разрезах:

- Изменение распределения портфеля по интервалам рейтинговой шкалы.
- Изменение распределения сегментов по интервалам рейтинговой шкалы.



При выявлении существенных изменений, рекомендуется более глубокий анализ.

Для оценки стабильности также используется индекс PSI, граничные значения для которого указаны в Приложении 1.

Стоит отметить, что показатель PSI, обладает существенным недостатком. Он не поддается строгой статистической трактовке, то есть устанавливаемые граничные значения являются строго экспертными, при этом нельзя оценить, насколько статистически достоверен полученный результат теста.

Для анализа стабильности и репрезентативности выборки возможно использование точных статистических тестов таких, как Хи-квадрат и тест Колмогорова-Смирнова на идентичность распределений величин в анализируемых выборках.

Анализ матриц миграций рейтингов проводят для выявления, насколько стабильны рейтинги отдельных клиентов. Банк самостоятельно определяет критерии по признанию рейтингов стабильными.

## 2.4. Концентрация

Тесты на концентрацию направлены на выявление избыточных концентраций клиентов в некоторых интервалах рейтинговой шкалы. Избыточные концентрации могут свидетельствовать о несоответствии используемой рейтинговой шкалы портфелю.

Концентрацию можно оценить с помощью **Индекса Херфиндаля  $HI$**  и **скорректированного Индекса Херфиндаля  $HI^{adj}$** :

$$HI = \sum_i \left(\frac{N_i}{N}\right)^2 \qquad HI^{adj} = \frac{\sum_i \left(\frac{N_i}{N}\right)^2 - \frac{1}{J}}{1 - \frac{1}{J}}$$

где  $i$  – индекс по интервалам рейтинговой шкалы,  $J$  – количество интервалов рейтинговой шкалы,  $N$  – общее количество наблюдений,  $N_i$  – количество наблюдений в интервале  $i$ .

В Приложении 1 приведены пороговые значения для оценки концентрации.

Анализ целесообразно проводить в следующих разрезах:

- На уровне портфеля в целом.
- На уровне сегментов.
- На уровне модулей (при наличии PD на уровне модулей).

Сама по себе высокая концентрация наблюдений в области низких (высоких) значений вероятности дефолта не может приводить к выводу о низком качестве работы модели, поскольку она может быть обусловлена тем, что в портфеле в определенный момент времени преобладают заемщики с низким (высоким) уровнем риска. Это определяется спецификой продукта и сегмента. Соответственно, с точки зрения валидации моделей также имеет смысл исследовать динамику показателей концентрации заемщиков в тех или иных диапазонах вероятности дефолта.

## 2.5. Влияние процесса на результаты работы модели

В рамках проведения ежегодной валидации необходимо оценить, каким образом влияет процесс присвоения рейтинга на качество работы модели. Рекомендуется проанализировать влияние:

- Замены отсутствующих и/или нерелевантных значений из источников при расчете рейтинга на значения по умолчанию.
- Ручных корректировок рейтинга.

Объем и перечень тестов для выполнения проверок, определяется банком самостоятельно.

## II. Методические рекомендации по валидации внутренних моделей оценки LGD и EAD

Под валидацией модели расчета кредитного риска понимается набор процессов и мероприятий, используемых для оценки корректности дифференциации риска рейтинговой моделью, а также для оценки адекватности характеристики компонент риска параметрами PD, LGD, EAD.

Несмотря на то, что на текущий момент существует достаточное количество способов количественной и качественной оценки моделей, Базельский комитет отмечает, что выделить универсальный подход валидации, эффективный для всех банков и портфелей, не представляется возможным.

Данный документ предполагает, что выбор подхода к валидации должен основываться на сложности модели и уровне риска портфеля банка. Чем сложнее применяемая модель, и чем больший риск несет банк, тем более глубоким и подробным должен быть процесс валидации.

В зависимости от сложности модели и степени риска портфеля банка, можно выделить следующие типы валидации:

- *Первичная валидация.* Применяется для простых моделей, используемых для оценки портфелей с низким риском.
- *Продвинутая валидация.* Используется для моделей средней сложности, или в случаях, когда портфель банка характеризуется средним уровнем риска.
- *Комплексная валидация.* Применяется для сложных моделей, или для ситуаций, когда банк несет существенный кредитный риск

В зависимости от типа, валидация включает в себя следующие этапы:

1. Оценка модели пользователем и степень интеграции в кредитный процесс
2. Количественное тестирование
  - Оценка эффективности модели с помощью статистических тестов (дискриминационная способность, качество калибровки и стабильность)
  - Бэнчмаркинг, бэктестинг
3. Качественное тестирование
  - Оценка дизайна/ методологии модели
  - Оценка процесса разработки модели
  - Оценка управления данными при разработке модели

Первичная валидация включает в себя оценку модели пользователем и степень интеграции в кредитный процесс. Продвинутая валидация, помимо первого этапа, включает в себя количественное тестирование модели. Комплексная валидация совмещает в себе все три этапа.

В рамках предложенного подхода, валидация LGD/EAD-моделей имеет следующие этапы (рис.1):

1. Оценка модели пользователем
2. Оценка результатов LGD/EAD-модели
3. Валидация методики расчета
4. Валидация входных данных и процесса формирования исходной выборки



Рис.1. Этапы валидации LGD/EAD-модели

Настоящий документ предоставляет описание особенностей этапов качественной и количественной валидации LGD/EAD-моделей, а также содержит информацию о применяемых статистических инструментах.

## 1. Первичная валидация LGD/EAD-моделей: Оценка модели пользователем

### 1.1. Восприятие модели

Оценка восприятия LGD/EAD-модели пользователем предполагает комплексный анализ мнений специалистов кредитных подразделений о точности, эффективности и полезности валидируемой модели. Целью данного этапа валидации является определение того, насколько результаты модели соответствуют ожиданиям конечных пользователей.

Анализ восприятия модели включает в себя следующие аспекты:

- *Точность.* Отражает частоту совпадения результатов модели с мнением кредитных специалистов и возможность отличить незначительные колебания в результатах модели от существенных изменений кредитного риска.
- *Полезность.* Отражает оценку пользователем вклада модели в процесс управления кредитными рисками.
- *Простота использования.* Чрезмерное техническое усложнение моделей приводит к низкому уровню восприятия и понимания результатов.
- *Соответствие решаемым задачам.* Используемая модель должна покрывать максимальное количество присутствующих в банке типов продуктов и групп заемщиков.

### 1.2. Использование модели и степень интеграции в кредитный процесс

Анализ степени использования и уровня интеграции в кредитный процесс предполагает оценку того, насколько активно банк применяет результаты работы модели для анализа рисков и принятия бизнес-решений.

Данный этап предполагает следующие области для валидации:

- Интеграция LGD/EAD-модели в следующие системы и процессы:
  - Процесс принятия решений по кредитной сделке
  - Система анализа кредитоспособности заемщика
  - Процесс мониторинга кредитного риска (Оценка портфеля в целом и каждого заемщика по отдельности)
- Использование результатов LGD/EAD-модели в следующих целях:
  - Расчет кредитного риска по портфелю в целом
  - Установление лимитов
  - Ценообразование продуктов
  - Расчет резервов
  - Расчет уровня достаточности капитала

### **1.3. Требования к документации**

Анализ документации модели предполагает оценку четкости, полноты и прозрачности ее описания. Данный этап валидации включает проверку следующих аспектов:

- Насколько подробно описана модель? Содержит ли документация описание следующих параметров:
  - Основные предположения
  - Ограничения и случаи применимости
  - Описание входных данных и их источников
  - Методология расчета
- Возможность воспроизведения результатов моделирования третьей стороной на основе документации.
- Возможность новых сотрудников использовать модель, на основе документации, в случае смены команды риск-менеджмента.

## **2. Продвинутая валидация LGD/EAD-моделей: Оценка результатов**

### **2.1. Оценка эффективности модели**

Количественная валидация моделей оценки LGD/EAD представляет собой оценку эффективности модели с помощью статистических тестов. В процессе количественной оценки эффективности системы расчета LGD/EAD происходит сравнение значений, предсказанных моделью с потерями от реально наблюдаемых дефолтов (бэктестинг), либо с оценками LGD/EAD, полученными при помощи других моделей (бэнчмаркинг).

Оценка эффективности должна включать определение следующих параметров модели:

- Стабильность
- Дискриминационная способность
- Качество калибровки

Стабильность характеризует способность системы адекватно моделировать причинно-следственную связь между факторами и оцениваемыми параметрами. Стабильная модель не чувствительна к ложным эмпирическим корреляциям, и качество ее прогнозов не меняется с течением времени.

Дискриминационная способность, в рамках валидации систем оценки параметров кредитного риска, характеризует способность модели различать платежеспособных заемщиков и тех, кому угрожает дефолт.

Качество калибровки характеризует степень точности, с которой система прогнозирует тот или иной параметр.

Для получения оценок стабильности, дискриминационной способности и качества калибровки моделей оценки LGD/EAD используются следующие статистические инструменты (табл.1)<sup>4</sup>.

Таблица 1

**Статистические инструменты валидации моделей LGD и EAD**

	LGD	EAD
Стабильность	Индекс стабильности системы / критерий Колмогорова-Смирнова	
Дискриминационная способность	Loss Capture Ratio /CLAR/Ранговая корреляция Спирмена	
Качество калибровки	Loss Shortfall/Среднее абсолютное отклонение/ Т-статистика	Т-статистика

Выбор конкретных инструментов валидации производится экспертом и основывается на особенностях конкретной модели. Методика использования инструментов, а также предлагаемый список применяемых тестов и их пороговых значений представлены в Приложении.

## 2.2. Бэнчмаркинг

Одним из требований БКБН является необходимость использования количественных инструментов для сравнения результатов модели с релевантными внешними источниками.<sup>5</sup>

В общем случае, под бэнчмаркингом понимается сравнение внутренних оценок параметров кредитного риска между различными банками или/и другими внешними источниками<sup>6</sup>. Также возможен вариант внутреннего бэнчмаркинга, который заключается в сравнении банком значений, полученных при помощи модели с оценками экспертов или результатами альтернативной внутренней модели, используемой для валидации действующей. Разновидностью внутреннего бэнчмаркинга можно считать подход Champion/Challenger: основная модель сравнивается с конкурирующей, после чего выбирается лучшая из двух представленных. Данный подход может использоваться в процессе разработки моделей.

Применение бэнчмаркинга представляет для банка альтернативную возможность проведения количественной оценки в условиях недостатка данных, что является особенно актуальным для оценки LGD и EAD. Являясь частью валидации, бэнчмаркинг делает данный процесс более гибким, так как банк самостоятельно выбирает источник внешних данных, который будет наиболее соответствовать портфелю банка. В этом отношении бэнчмаркинг

<sup>4</sup> Backtesting Framework for PD, EAD and LGD, Master Thesis, Bauke Maarse

<sup>5</sup> International Convergence of Capital Measurement and Capital Standards, BCBS, п.502

<sup>6</sup> Studies on the Validation of Internal Rating Systems, BCBS (2005) с. 9

позволяет расширить формальный подход к валидации (проведение которого не всегда возможно) эмпирической оценкой<sup>7</sup>.

Источниками данных для бенчмаркинга могут быть кредитные бюро, рейтинговые агентства, пулы данных или экспертные оценки.

Выбор бенчмарка должен основываться на характеристиках исследуемой модели.<sup>8</sup> При выборе внешнего источника банку необходимо учитывать следующие особенности:

- Качество внешних данных
- Различные методологии и процессы, на которых основаны сравниваемые модели
- Различное определение дефолта и потерь
- Правовые ограничения
- Возможно, различный подход к расчету LGD

В то же время, проведение бенчмаркинга с использованием внешних источников может быть затруднительным на практике вследствие:

- отсутствия внешних моделей для бенчмаркинга по значимой части заемщиков в портфеле банка;
- нерепрезентативности анализируемой выборки в силу специфики сбора данных;
- отсутствия учета ряда специфических для данного кредита условий (процентная ставка, срок кредита, наличие и стоимость залога и т.п.) во внешних источниках.

Для интерпретации результатов бенчмаркинга строится матрица соответствий.

Выбирается множество заемщиков, для которых определяются оценки LGD по внутренней модели банка и по бенчмарку. Оба полученных множества оценок LGD упорядочиваются по возрастанию и разбиваются на N диапазонов каждое. Строится матрица соответствия, которая содержит N строк и N столбцов. Каждая ячейка на пересечении строки i и столбца j содержит значение  $p_{ij}$  - долю элементов выборки (или долю суммарной EAD), которая во внутренней модели относится к i-ому диапазону, а во бенчмарку – к j-ому. Таким образом, диагональная матрица представляет собой идеальное соответствие моделей.

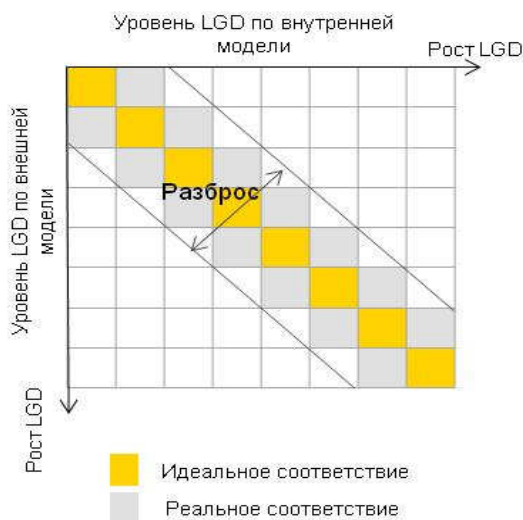


Рис.2. Матрица соответствий

7 Studies on the Validation of Internal Rating Systems, BCBS (2005) с.102

8 Studies on the Validation of Internal Rating Systems, BCBS (2005), с.98

Для оценки степени изменений уровней LGD могут использоваться индикаторы перехода и также индикатор отображения (mapping indicator):

Расстояние в количестве уровней, взвешенное по числу наблюдений:

$$I_{Trans}^A(P_{N,N}) = \sum_{i=1}^N \sum_{j=1}^N \frac{(i-j)^2}{(N-1)^2} \times \frac{Numbers_i}{\sum_{k=1}^N Numbers_k} \times p_{ij}$$

Расстояние в количестве уровней, взвешенное по величине подверженности кредитному риску:

$$I_{Trans}^A(P_{N,N}) = \sum_{i=1}^N \sum_{j=1}^N \frac{(i-j)^2}{(N-1)^2} \times \frac{Exposure_i}{\sum_{k=1}^N Exposure_k} \times p_{ij}$$

Индикатор отображения:

$$I_{MAP}(P_{N,N}) = \sum_{i=1}^N \sum_{j=1}^N \frac{(i-j)^2}{(N-1)^2} \times p_{ij}$$

Помимо бенчмаркинга, данные показатели могут применяться к миграционным матрицам и использоваться, как инструмент оценки степени изменения прогнозного показателя LGD за определенный отрезок времени (например, 1 год).

В процессе бенчмаркинга должна анализироваться динамика изменения матрицы соответствий. Если предположить, что на рис.7 изображена матрица в году N, то возможные изменения вида матрицы соответствий представлены на рис.8.

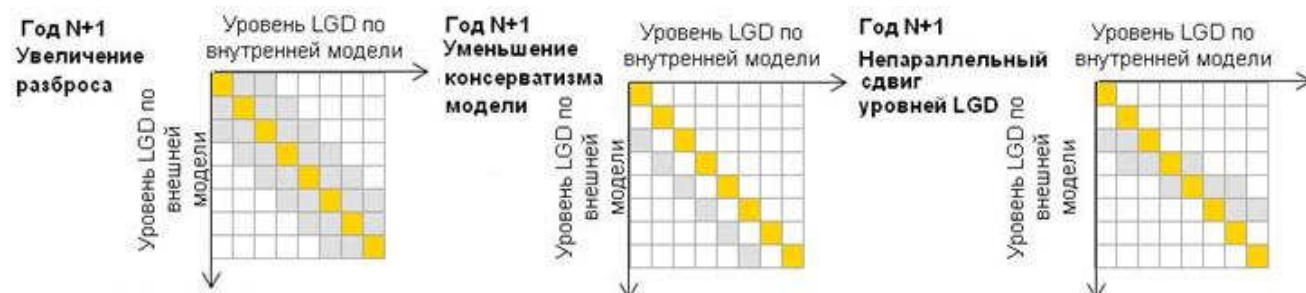


Рис.3. Возможные изменения матрицы соответствий

Увеличение разброса по отношению к бэнчмарку может свидетельствовать об уменьшении точности модели. Сдвиг диагонали матрицы соответствий может свидетельствовать об уменьшении степени консерватизма модели по сравнению с бэнчмарком. Непараллельный сдвиг в матрице соответствий может свидетельствовать о завышении оценки LGD для одних групп заемщиков и занижении оценки для других групп.

Бенчмаркинг должен использоваться как дополнение, но не как замена статистических методов валидации.

### 2.3. Построение тестовой выборки

В случае, когда валидация является одним из этапов разработки модели, вся исходная выборка делится на обучающую (примерно 70% данных) и тестовую (примерно 30% данных) выборки. В случае наличия большего количества данных по сравнению с количеством, требуемым для разработки модели, тестовая выборка должна быть расширена.

При наличии необходимого набора данных, может применяться перекрестная валидация. В этом случае тестовая выборка делится на k групп данных (например, на 10 равных групп



данных) и используется в целях валидации модели в количестве  $k$  раз. Результат валидации представляется в виде среднего показателя по каждому набору данных

В случае недостатка необходимых данных в тестовой выборке, может применяться подход статистического бутстреппинга. Результаты валидации в случае использования статистического бутстрэппинга должны быть интерпретированы с учетом всей предоставленной информации.



Рис.4. Подходы к формированию тестовой выборки

В рамках валидации моделей LGD/EAD возможно применение двух подходов к разделению исходной выборки на обучающую (серые маркеры) и тестовую (оранжевые маркеры) (см. рис.Рис.4. Подходы к формированию тестовой выборки). Горизонтальная ось каждого графика обозначает зависимость данных от времени наблюдения.

#### 1) Out of Sample.

Данные для тестовой выборки, не участвующие в процессе обучения модели, выбираются из исходной выборки случайным образом. При этом не накладывается никаких ограничений на время дефолта и тип подвыборки. Единственным условием является то, что обучающая и тестовая выборки не пересекаются. Подход Out of Sample предполагает, что свойства данных остаются стабильными на протяжении всего времени (процесс является стационарным). Так как тестовая выборка формируется случайным образом, то валидация происходит с учетом свойств всех типов подвыборок, таким образом, сохраняется первоначальный тип распределения LGD/EAD.

#### 2) Out of Sample, Out of Time.

Для обучающей выборки выбираются данные, предшествующие определенной дате в прошлом, в то время как тестовая выборка формируется за счет данных после выбранной даты. Так как валидация происходит с использованием данных из другого временного периода, то рассматриваемый метод позволяет установить наличие временной зависимости. Так как при формировании тестовой выборки не накладывается ограничений на типы используемых подвыборок, то, как и в первом случае, сохраняется первоначальный вид распределения LGD/EAD.

Наименее строгим с точки зрения валидации является первый подход. Выбор конкретного метода формирования тестовой выборки для валидации должен проводиться экспертом и отвечать особенностям валидируемой модели.

### 3. Комплексная валидация LGD/EAD-моделей: Качественное тестирование методики

#### 3.1. Анализ драйверов риска и особенностей расчета прогнозного значения LGD

При качественной валидации EAD/LGD-моделей должен проводиться анализ предположений о драйверах риска, которые используются в модели.



Для LGD-моделей можно выделить следующие драйверы риска (не ограничиваясь представленным списком):

### **3.1.1. Макроэкономические факторы**

В зависимости от особенностей портфеля и предполагаемого источника возмещения потерь в случае дефолта, для получения оценок LGD банк может использовать различные макроэкономические показатели: рост ВВП, уровень безработицы, уровень процентных ставок, курсы валют, ценовые индексы и т.п.

Одним из требований базельского комитета для LGD-моделей<sup>9</sup> является учет условий экономического спада для получения оценки LGD. Нисходящая фаза экономического цикла характеризуется повышенной частотой дефолтов и относительно более высокими уровнями LGD. Таким образом, качественная валидация LGD-моделей должна включать анализ того, насколько полно учитываемые макроэкономические драйверы отражают макроэкономическую ситуацию и, как следствие, насколько консервативными являются результаты модели с учетом нисходящей фазы экономического цикла.

### **3.1.2. Параметры заемщика и характеристики кредитного продукта**

В процессе валидации необходимо проанализировать, каким образом модель учитывает параметры заемщика и характеристики кредитного продукта:

Для заемщика:

- тип (суверенный, юридическое лицо, малый бизнес и т.д.),
- страна или регион
- отрасль
- организационная структура
- и т.д.

Для кредитного продукта:

- тип (кредит, долговая бумага..)
- цель предоставления
- обеспеченность
- субординированность
- и т.д.

Данные характеристики могут являться существенными факторами, влияющими на процесс взыскания задолженности и, как следствие, на оценку LGD.

### **3.1.3. Параметры обеспечения**

Поступления средств от обеспечения зависят от возможности реализации актива, лежащего в его основе. Размер поступлений зависит не только от балансовой стоимости, но может определяться как особенностями конкретного актива (например, техническим состоянием), так и глубиной и активностью соответствующего рынка. Процесс качественной валидации должен затрагивать анализ того, как данные факторы учитываются моделью и насколько консервативными являются дисконты, применяемые к обеспечению в свете возможного развития стрессовой ситуации на соответствующем рынке.

---

<sup>9</sup> International Convergence of Capital Measurement and Capital Standards, BCBS, п.468

### **3.1.4. Гарантии**

Дополнительным фактором, влияющим на объем возможных поступлений в случае дефолта заемщика, являются параметры предоставленных гарантий. Такими параметрами являются финансовое положение поручителя, а также конкретные условия договора о гарантии.

### **3.1.5. Внутренние факторы банка**

Качественная валидация LGD-моделей должна включать анализ того, насколько особенности внутренних процессов банка учитываются при расчете ожидаемых потерь в случае дефолта. К наиболее важным, с данной точки зрения, процессам можно отнести подходы к взысканию просроченной задолженности, особенности мониторинга портфеля, частоту переоценки обеспечения и т.д.

Помимо анализа вышеперечисленных драйверов риска, при валидации моделей оценки LGD должны быть проанализированы следующие аспекты:

- Прослеживается ли корреляция между PD и LGD. Если да, то учитывается ли данная зависимость при расчете прогнозного LGD
- Соответствует ли определение дефолта, используемое для прогноза показателя LGD тому, что используется для прогноза PD

## **3.2. Анализ особенностей расчета реализовавшихся потерь**

Вследствие того, что процесс взыскания задолженности с заемщика, находящегося в состоянии дефолта, является длительным, а также из-за возможных особенностей организации процесса взыскания, в целях валидации модели оценки LGD необходимо проанализировать следующие вопросы:

- Каким образом в модели учитываются дефолты, по которым процесс взыскания задолженности еще не завершен
- Как определяется максимально возможный срок взыскания задолженности, после которого процесс считается завершенным
- Как учитываются взыскания не денежного характера

В процессе валидации должно быть получено объективное обоснование применения того или иного подхода к решению вышеуказанных вопросов.

## **3.3. Анализ драйверов риска и особенностей расчета прогнозного значения EAD**

Так как единственным случайным фактором для расчета EAD является CCF, то величина EAD во многом определяется характеристиками кредитного продукта и типичным поведением заемщиков. Таким образом, в процессе валидации необходимо проанализировать, насколько данные аспекты учитываются в действующей модели, а именно:

- Волатильность открытых кредитных линий
- Типичное поведение заемщика
- Характер ставки кредитования. Для кредитных продуктов с плавающей ставкой, оценка EAD зависит от значений и волатильности соответствующего бенчмарка (LIBOR итп.)
- Возобновляемость кредитной линии
- Дополнительные условия договора. Возможно наличие корреляции значений EAD и кредитного рейтинга заемщика в момент дефолта. Данная корреляция может быть следствием особенностей условий кредитного договора, которые ограничивают размер возможных списаний по кредитным линиям в случае снижения рейтинга клиента.

• Реструктуризация. В случае если заемщик испытывает финансовые затруднения, либо находится в состоянии дефолта, реструктуризация задолженности может привести к

установлению более строгих условий договора, и снизить вероятность списания клиентом неиспользованной части кредитной линии.

- Время до окончания контракта. Чем больше времени остается до окончания договора, тем больше вероятность, что финансовое положение заемщика ухудшится и, как следствие, ему потребуется списать неиспользованную часть кредитной линии.

- Альтернативные источники финансирования. Чем больше у заемщика альтернативных источников финансирования, тем меньше ожидаемое значение EAD.

- Текущий объем использованной кредитной линии.
- Насколько оправдана принятая сегментация задолженности для оценки EAD
- Возможное наличие корреляции между PD и EAD

#### **4. Комплексная валидация LGD/EAD-моделей: Качественное тестирование данных**

В ходе валидации необходимо установить, отвечают ли используемые данные общим требованиям (точность, объективность, достоверность, полнота, независимость, последовательность), а также требованиям, обусловленным спецификой расчета LGD/EAD.

##### **4.1. Анализ методов агрегации и построения исходной выборки**

В случае, если сопоставимость данных достигается путем их частичного преобразования, то в рамках валидации модели необходимо проанализировать, насколько обоснованным является применение тех или иных методов корректировки.

Кроме того, если для построения исходной выборки применяются какие-либо способы масштабирования или фильтрации данных, необходимо проанализировать, какое влияние используемые методы оказывают на свойства распределения.

В частности, для валидации модели оценки LGD, следует учесть, каким образом обрабатываются случаи, когда реализованное значение LGD меньше или равно 0. Если такие наблюдения исключаются из исходной выборки, необходимо проанализировать, насколько в данной модели совпадают определения дефолтов для PD и LGD.

##### **4.2. Проверка сопоставимости данных**

Под сопоставимостью данных понимается единство методологии расчёта, единиц измерения, полноты охвата наблюдением явления, а также соблюдение других условий сравнимости показателей.

В случае если в качестве входных данных для модели оценки LGD используются исторические данные за продолжительный отрезок времени, необходимо проанализировать, каким образом достигается их сопоставимость между собой, а именно:

- Как согласуются различные определения дефолта в разные периоды времени
- Как обрабатываются случаи, когда заемщик из состояния «дефолт» вернулся в состояние «нет дефолта»
- Как учитываются изменения в методиках и стоимости взыскания просроченной задолженности, определения ставок дисконтирования и т.д.

Вследствие ограниченной доступности истории по дефолтам, выборка для определения LGD может формироваться, в том числе, из внешних источников данных. В рамках валидации необходимо проанализировать, каким образом в ходе построения исходной выборки достигается сопоставимость используемых внешних и внутренних данных, а именно:

- Как учитываются различия в подходах к взысканию просроченной задолженности
- Как учитывается разная стоимость взыскания просроченной задолженности
- Как учитываются разные ставки дисконтирования при расчете Recovery Rate
- Как согласуются определения дефолта во внешних и внутренних данных

- Насколько качественны внешние данные

### **4.3. Наличие в выборке данных, соответствующих экономическому спаду**

Построение точной модели оценки LGD требует включения в исходную выборку данных, соответствующих нисходящей фазе экономического цикла, которая характеризуется повышенной частотой дефолтов и относительно более высокими уровнями LGD<sup>10</sup>. Таким образом, в рамках валидации, необходимо проанализировать, включаются ли в исходную выборку такие периоды.

---

<sup>10</sup> Studies on the Validation of Internal Rating Systems, Basel Committee on Banking Supervision (2005)

### **III. Методические рекомендации по валидации процесса присвоения рейтинга, использования риск-параметров и адекватности систем контроля**

#### **A. Валидация процесса присвоения рейтинга, использования риск-параметров и адекватности систем контроля**

В соответствии с Международной конвергенцией измерения капитала и стандартов капитала (Базель 2) термин «рейтинговая система» включает в себя все методы, процедуры, системы контроля и сбора данных, а также информационные системы, которые используются при оценке кредитного риска, присвоению внутренних рейтингов и количественным оценкам вероятности дефолта и убытка. Таким образом, валидация рейтингового процесса подразумевает оценку таких важных аспектов как: корректность рассчитанного рейтинга, использование риск-параметров, адекватность процедур контроля. Подобная оценка носит скорее качественный характер и заключается не в простой проверке соответствия нормативным требованиям, но и по возможности, анализе применяемой практики (также в сравнении с нормативными требованиями). Однако нужно обратить внимание на то, что со временем «строгость» метрик по портфелю должна основываться на экспертной оценке силы корреляции количественных характеристик, получаемой путем соотнесения (корреляции) их со случившимися дефолтами в рамках жизненных циклов кредитов отдельных заемщиков и роста всего портфеля.

Данный раздел документа содержит рекомендации по оценке процесса присвоения рейтинга, внутреннего использования риск-параметров, системы контроля и степени одобрения системы участниками процесса и является приложением к Методическим рекомендациям по реализации подхода к расчету кредитного риска на основе внутренних рейтингов банков.

Группы тестов для проведения валидации:

1. Процесс присвоения рейтингов
  - 1.1. Внутренние нормативные документы
  - 1.2. Целостность
  - 1.3. Однородность и уникальность
  - 1.4. Обновление рейтинга и покрытие
  - 1.5. Обучение
2. Внутреннее использование
  - 2.1. Оценка степени использования риск-параметров
  - 2.2. Отчетность и мониторинг
3. Система контроля
  - 3.1. Сегментация и ведение групп заемщиков
  - 3.2. Качество данных, используемых для расчета рейтинга
4. Одобрение системы пользователями

Ответы на тестовые вопросы целесообразно структурировать следующим образом:

- Описание ситуации «как есть»: описание текущей ситуации в соответствии с имеющимися документами и опроса ответственных лиц. Описание должно быть исчерпывающим и понятным третьим лицам (ответы вида «да»/ «нет» недостаточны).

- Уровень адекватности: оценка адекватности в соответствии с «принципом светофора»: «зеленый» - полное выполнение всех нормативных требований, «желтый» - существуют некоторые отклонения/ недостатки, «красный» может означать наличие не только некоторых недостатков, но и некорректное использование рейтинговой системы. Тем не менее, «желтый» и «красный» результат выполнения теста может приводить к рекомендациям

как высокого, так и среднего и низкого приоритета в зависимости от степени существенности выявленных недостатков.

- Анализ недостатков: исчерпывающее и подробное описание возможных недостатков в зависимости от требований теста.
- Рекомендация: предложение определенных действий, позволяющих устранить выявленные недостатки. Степень значимости (приоритет) рекомендации также должен быть указан.
- Список документов: перечень документов, задействованных в подготовке ответа на тестовый вопрос.

Раздел тестов №4 «Одобрение системы пользователями» не предполагает оценки уровня адекватности и подготовки рекомендаций. Этот раздел, по сути, дает возможность собрать отзывы и впечатления пользователей от пользования рейтинговой системой.

Все группы тестов рекомендуется выполнять как минимум ежегодно, по результатам валидации должен быть представлен соответствующий отчет. Выполнение рекомендаций должно быть закреплено в плане действий с четким указанием ответственных лиц и сроков исполнения. Статус выполнения рекомендаций также оценивается в рамках ежегодной валидации.

## **1. Процесс присвоения рейтингов**

В рамках этой группы тестов предлагается ответить на следующие вопросы:

### **1.1. Внутренние нормативные документы**

1. Существует ли внутренняя инструкция/ регламент, где подробно описаны роли и зоны ответственности подразделений, вовлеченных в рейтинговый процесс (включая ручные корректировки рейтингов)? Если да, отображены ли роли, зоны ответственности и иерархия, в которой четко определены полномочия по корректировке рейтингового процесса и инструментов в соответствии с лучшими практиками, в организационной структуре?

2. Соответствует ли фактический процесс присвоения рейтингов внутренней инструкции/ регламенту? При наличии расхождений, указать причины.

3. Имеются ли операционные инструкции по рейтинговой системе, включающие как минимум:

- Правила сегментации
- Факторы и компоненты, влияющие на вычисленный моделью результат
- Качественные вопросы
- Список возможных причин внесения ручных корректировок рейтинга

4. Каким образом регулируются случаи одновременного внесения различных корректировок рейтинга (ручная корректировка, ограничение по сроку действия рейтинга, влияние группы и т.д.)

5. Производится ли своевременное обновление этих документов ответственными подразделениями и последующее их предоставление заинтересованным лицам?

6. Существуют ли внутренние операционные инструкции по своевременному выявлению так называемых «технических» дефолтов (т.е. произошедшим не по причине ухудшения кредитоспособности)? При идентификации «ложноположительных» или «технических» дефолтов, рекомендуется их документирование, а также создание стандартной процедуры отслеживания изменений определения дефолта для минимизации случаев «ложноположительных» или «технических» дефолтов. Данные по рейтингу «ложноположительных» дефолтов должны использоваться в дальнейшем при перенастройке рейтинговых инструментов.

7. Каким образом до сотрудников, задействованных в рейтинговом процессе, доносится информация об обновлении модели, внедрении новой ИТ функциональности?

## **1.2. Целостность**

8. Имеется ли у сотрудников, осуществляющих утверждение рейтинга, заинтересованность в объемах кредитования и выручке? Уполномочены ли они принимать решение о выдаче кредита? Если да, то какие меры предосторожности обеспечивают независимость оценки? Отсутствует ли у клиентских менеджеров доступ к количественным и качественным характеристикам рейтинговых инструментов, индикаторам риска и их весу при принятии решения о выдаче/невыдаче?

## **1.3. Однородность и уникальность**

9. Существуют ли технические инструкции по последовательности/ комбинации неавтоматических компонентов (качественный вопросник, ручные корректировки рейтинга) для обеспечения одинаковой процедуры расчета всеми пользователями?

## **1.4. Обновление рейтинга и покрытие**

10. Соответствует ли процесс обновления рейтинга нормативным документам?

11. Применяются ли в банке правила устаревания рейтинга? Какова доля устаревших рейтингов на данный момент? Принимаются ли меры по снижению доли таких рейтингов?

12. Предусмотрена ли в банке возможность ручных корректировок рейтинга? Каков процент таких корректировок? Как он изменяется со временем?

13. Каждый ли клиент в портфеле получает рейтинговый класс в соответствии с рассчитанным PD? Каково текущее покрытие портфеля рейтингами и как оно изменяется со временем?

14. Присвоены ли каждой кредитной сделке в портфеле соответствующие параметры LGD и EAD (если применимо)? Происходит ли пересмотр кредитного рейтинга при изменении (увеличении или уменьшении) общей суммы обязательств или добавлении новых кредитных продуктов?

## **1.5. Обучение**

15. Проходили ли вовлеченные в рейтинговый процесс сотрудники соответствующее обучение, касающееся как с методологических, так и операционных аспектов рейтинговой системы? Проводится ли обучение аккредитованной организацией или другим (высшим) учебным заведением? Включает ли обучение получение сотрудниками банка сертификатов, позволяющих проводить обучение других сотрудников?

## **2. Внутреннее использование**

### **2.1. Степень использования риск-параметров**

16. Какова роль параметров оценки рисков (PD, LGD, EAD) в выдаче кредитов/ продлении сроков кредитования?

17. Влияют ли параметры оценки рисков на принятие решения ответственным за выдачу кредита лицом/ органом?

18. Используются ли внутренние параметры рисков в стратегии управления кредитным портфелем?

19. Отличаются ли оценки, используемые для управленческих/ целей бухгалтерского учета<sup>11</sup> от нормативных? Если «да», то подтверждено ли это документами, в которых четко отражено обоснование применения различных оценок и их связь с нормативными оценками?

## **2.2. Отчетность и мониторинг**

20. Отражено ли в существующей регулярной отчетности по портфелю:

- Параметры оценки рисков (PD, LGD, EAD) и их изменение со временем.
- Распределение портфеля по рейтинговой шкале и значениям LGD в том числе в сравнении с результатами предыдущего года.

21. Кто является получателем этой отчетности и с какой периодичностью? Предоставляется ли она высшему руководству?

22. Существует ли последующий внутренний мониторинг рейтингов (включая статистику по ручным корректировкам, устареванию рейтингов)? Кто является получателем этих отчетов? Существует ли процедура эскалации на случай критических результатов?

23. Осуществляется ли поиск проблемных мест в рейтинговом процессе, и отслеживается ли процесс их устранения? Информировается ли высшее руководство о текущем статусе?

24. Существует ли процесс мониторинга изменений рейтингов и кредитоспособности заемщиков/ портфеля во времени? Информировается ли высшее руководство о результатах?

## **3. Система контроля**

### **3.1. Сегментация и ведение групп взаимосвязанных заемщиков**

25. Каким образом обеспечивается корректность сегментации и применяемой рейтинговой модели?

26. Каким образом подтверждается корректность состава и связей в группах взаимосвязанных заемщиков и их рейтингов? Есть ли «единый профиль клиента», покрывающий группы связанных компаний, акционеров и их частные активы и обязательства во всех сегментах (напр., розница, СМБ, private banking, корпоратив) и позволяющий таким образом определить отдельные объемы обязательств и общую сумму обязательств?

27. Существуют ли автоматические проверки корректности сегментации и установления экономических/ юридических связей в группах взаимосвязанных заемщиков?

28. Были ли обнаружены какие-нибудь отклонения? Если «да», то есть ли понимание путей решения?

### **3.2. Качество данных, используемых для расчета рейтинга**

29. Существует ли актуальная внутренняя политика по качеству данных, действующая в отношении параметров кредитного риска?

30. Имеется ли соответствующий процесс сертификации данных в случае ручного ввода информации для расчета рейтинга?

---

<sup>11</sup> Резервирование, ценообразование, оценка эффективности, принятие решений и т.д.



31. Существует ли минимум два уровня контроля за корректностью вводимых данных для расчета рейтинга?

32. Существует ли шаблон для отражения результатов процедур контроля качества данных?

33. Внедрен ли процесс по контролю и управлению качеством данных, в том числе соответствующая процедура эскалации?

Существуют ли процедуры для 1) обнаружения (временных) проблем качества данных (напр., выпадающих значений) в рамках анализа данных и поддержки принятия решений, 2) корректировки обнаруженных проблем с данными напрямую в соответствующих отчетах и аналитике и 3) исправления данных в исходных системах?

#### **4. Одобрение системы пользователями<sup>12</sup>**

34. Понятен ли пользователям рейтинговый процесс (т.е. роли, зоны ответственности)?

35. Являются ли правила сегментации однозначными и ясными?

36. Являются ли входные переменные модели (PD и LGD) ясными и прозрачными?

37. Все ли важные факторы риска включены в модель (PD и LGD)?

38. Понятны ли качественные вопросы? Являются ли они однозначными? Существует ли процедура регулярного/ad hoc регрессивного анализа, проводимого для определения корреляции качественных и количественных показателей модели и имевших место дефолтов (PD)?

## **В. Валидация процесса по снижению кредитного риска**

Этот раздел документа содержит руководство по анализу и валидации процессов оформления, оценки и поддержке гарантий/обеспечения, приемлемых для целей снижения кредитного риска (CRM), в рейтинговой системе.

Валидация CRM процесса является в основном качественной и заключается в проверке соответствия нормативным требованиям и, по возможности, анализе применяемой практики (также в сравнении с нормативными требованиями).

Группы тестов, рекомендуемые для выполнения в рамках валидации CRM процесса:

1. Юридическая однозначность.
2. Своевременная ликвидация.
3. Организационные требования.
  - 3.1. Общие требования.
    - 3.1.1. Политики и процедуры.
    - 3.1.2. Оценка и процесс мониторинга.
    - 3.1.3. Управление данными.
    - 3.1.4. Контроль.
    - 3.1.5. Отчетность.
    - 3.1.6. Организационная структура.
  - 3.2. Специальные требования.

---

<sup>12</sup> Для полностью автоматизированного расчета рейтинга данный раздел не заполняется.

3.2.1. Фондированное кредитное обеспечение.

3.2.2. Кредитное обеспечение в виде гарантий/поручительств.

4.1. Раскрытие информации.

Ответы на тестовые вопросы целесообразно структурировать следующим образом:

- Описание ситуации «как есть»: описание текущей ситуации в соответствии с имеющимися документами и опроса ответственных лиц. Описание должно быть исчерпывающим и понятным третьим лицам (ответы вида «да»/ «нет» недостаточны).

- Уровень адекватности: оценка адекватности в соответствии с «принципом светофора»: «зеленый» - полное выполнение всех нормативных требований, «желтый» - существуют некоторые отклонения/ недостатки, «красный» может означать наличие не только некоторых недостатков, но и некорректное использование рейтинговой системы. Тем не менее, «желтый» и «красный» результат выполнения теста может приводить к рекомендациям как высокого, так и среднего и низкого приоритета в зависимости от степени существенности выявленных недостатков.

- Анализ недостатков: исчерпывающее и подробное описание возможных недостатков в зависимости от требований теста.

- Рекомендация: предложение определенных действий, позволяющих исправить выявленные недостатки. Степень значимости (приоритет) рекомендации также должен быть указан. Также должен применяться критерий материальности на основе определения возможных последствий выявленных недостатков и приоритизации их устранения.

- Список документов: перечень документов, задействованных в подготовке ответа на тестовый вопрос.

Все группы тестов рекомендуется выполнять ежегодно, по результатам валидации должен быть представлен соответствующий отчет. Выполнение рекомендаций должно быть закреплено в плане действий с четким указанием ответственных лиц и сроков исполнения. Статус выполнения рекомендаций также оценивается в рамках ежегодной валидации.

## **1. Юридическая однозначность**

1. Требуется ли политика банка проверки юридической силы, действительности, обязательного характера и возможности принудительного исполнения мер по обеспечению защиты кредита во всех соответствующих юрисдикциях, в том числе в случаях неплатежеспособности и банкротства заемщика и/или лица, предоставившего обеспечение?

2. Осуществляет ли банк следующие процедуры контроля:

- Полное и свободное право банка на принудительную реализацию обеспечения в судебном порядке? Проводит ли банк исследования соответствующей судебной практики, с целью пересмотра ценности некоторых залогов с точки зрения возможности взыскания и реализации в допустимые сроки? Если да, производит ли банк регулярный пересмотр приемлемости залога на основе возможности взыскания и реализации в допустимые сроки и/или пересмотра займа с учетом факторов риска (“haircuts”)?

- Были ли приняты все необходимые меры по обеспечению юридической силы, действительности, обязательного характера и возможности принудительной реализации обеспечения по кредиту в соответствии с действующим законодательством (включая получение и хранение соответствующей документации, однозначно устанавливающей наличие обеспечения по кредиту)?

- Лицо, предоставившее обеспечение, в соответствии с действующим законодательством, не имеет права ходатайствовать о защите, которая может лишить представленное обеспечение юридической силы, действительности, обязательного характера и возможности наложения принудительного взыскания (исключая возможность подачи иска об аннулировании)?

3. Внедрил ли банк процедуру пересмотра юридической силы документов (шаблонов документов), подтверждающих предоставление гарантии и залогового обеспечения, в случае изменения действующего законодательства?

4. Хранит ли банк всю документацию по обеспечению кредита в надежном и безопасном месте? Существуют ли меры по отслеживанию доступа сотрудников к документации (ИТ инструменты или вручную)?

## **2. Своевременная ликвидация**

5. Внедрены ли в Банке процедуры и технологии принятия быстрых мер по реализации активов, обеспечивающих требование (через ликвидацию активов и получение дохода от реализации, или посредством прямого приобретения актива, заложенного в качестве обеспечения, например, передачи прав на недвижимость)? Производит ли банк регулярные проверки того, что залог по-прежнему является собственностью заемщика? Существуют ли в банке временные рамки реализации залогов/гарантий, в которых минимизация LGD экономически обоснована?

## **3. Организационные требования**

### **3.1. Общие требования**

#### **3.1.1. Политики и процедуры**

6. Документировал ли и оформил ли банк политики и процедуры, касающиеся типов инструментов CRM, использующихся в нормативных расчетах, их количество и взаимосвязь с управлением совокупным профилем риска?

7. Соответствуют ли указанные политики и процедуры нормативным документам?

#### **3.1.2. Оценка и процесс мониторинга**

8. Имеются ли у банка документированные и формализованные политики и процедуры, касающиеся оценки и мониторинга инструментов CRM?

9. Использует ли банк подходы к оценке, которые одинаковым образом применяются ко всем видам залогового обеспечения и гарантиям, относящимся к одной категории?

10. Адекватно ли отслеживается стоимость залога, особенно когда рынок подвержен значительным изменениям условий? Существуют ли основанные на макроэкономических показателях и показателях портфеля триггеры проведения ситуативной переоценки залогов на уровне клиента или портфеля? Будут ли проведены последовательные изменения в процедурах предоставления гарантий и пересмотра рейтинга клиентов в случае значительных изменений оценки залогов?

11. Имеются ли у банка установленные процедуры и механизмы контроля, управления и, при необходимости, оценки риска несостоятельности, снижения или прекращения действия обеспечения («остаточные риски»)?

12. Имеются ли у банка необходимые процедуры и механизмы контроля, управления и, при необходимости, оценки потенциального риска концентрации в отношении определенных контрагентов или видов обеспечения?

13. Проводит ли банк проверку на предмет отсутствия существенной взаимосвязи между полученным заложенным обеспечением или полученной гарантией и кредитоспособностью заемщика или других связанных лиц, как на начальном этапе, так и на постоянной основе?

### **3.1.3. Управление данными**

14. Должным ли образом архивируются данные об обеспечении и гарантиях с целью возможности ретроспективного восстановления информации?

15. Требуют ли политики банка при регистрации залогов и гарантий ввода всей информации, необходимой для их обработки, в информационную банковскую систему (для оценки, мониторинга, прогноза LGD) в рамках проведения процедур, установленных банком?

16. Существует ли процесс регистрации новых видов обеспечения и их соответствующего связывания с категориями, определенными для целей оценки CRM и LGD?

### **3.1.4. Контроль**

17. Существует ли независимый контроль для проверки правильной «привязки» каждой гарантии/ обеспечения к конкретному кредиту, в обеспечение которого они предоставлены?

18. Существует ли независимый контроль для проверки правильности связи залогового обеспечения и гарантий с категориями, определенными для целей оценки CRM и LGD?

19. Внедрил ли банк процедуру проверки правильности введенной в систему суммы, срока действия и другой важной информации, касающейся идентификации и оценки обеспечения по кредиту?

20. Оформлены ли должным образом виды контроля и связанные с ними обязанности, и существует ли этому документальное подтверждение?

### **3.1.5. Отчетность**

21. Как часто предоставляется отчетность высшему руководству в отношении покрытия риска, выраженного в обеспечении и гарантиях?

22. Содержат ли отчеты информацию о виде обеспечения и гарантий, а также информацию об их стоимости и наличии каких-либо концентраций?

23. Подлежит ли занесению в отчетность результаты контроля, осуществляемого в отношении CRM?

### **3.1.6. Организационная структура**

24. В случае если обязанность по проверке юридической однозначности не делегирована централизованной структуре, принимает ли банк специальные организационные меры по обеспечению универсальности оценки внутренних структур и рабочих процессов, например, через выпуск соответствующих указаний и использование стандартизированных договоров?

25. Являются ли службы, ответственные за юридический контроль независимыми от подразделений/ сотрудников, заключающих договора?

26. Являются ли службы, ответственные за внесение гарантий и обеспечение в системы банка, независимыми от тех, что несут ответственность за управление бизнесом?

27. Должным ли образом оформлены эти обязанности?

## **3.2. Специальные требования**

### **3.2.1. Фондированное кредитное обеспечение**

28. В случае наличия материального или финансового обеспечения, каким образом банк подтверждает, что оценка обеспечения не превышает рыночную стоимость (например,

посредством обращения к независимым внешним оценщикам или получения котировок официальных рынков)?

29. В случае наличия материального обеспечения, отличного от финансового или залога недвижимости, каким образом банк проверяет, что стоимости, полученные посредством оценочных механизмов или источников котировок, незначительно отличаются от чистой суммы, которую банк получит при реализации обеспечения?

### **3.3. Кредитное обеспечение в виде гарантий/поручительств**

30. Отражают ли кредитные политики банка критерии приемлемости различных видов поручителей и требуют ли применения правил рейтингования, предусмотренных для расчета PD заемщика?

31. В случае условных гарантий (при которых гарант может не исполнять свои обязательства при выполнении определенных условий), имеются ли у банка четко описанные критерии для корректировки параметров риска в зависимости от способности гаранта выполнять свои обязательства по гарантиям, вероятного графика поступления платежей, степени, в которой способность гаранта исполнять свои обязательства по гарантиям связана с платежеспособностью заемщика и пределов остаточного риска по заемщику? В случае, если клиент нарушит один или несколько ковенантов договора в течение срока действия кредитного соглашения, есть ли возможность принятия адекватных контр-мер, включающих (но не ограничивающихся) 1) снижение рейтинга клиента, 2) снижение объема внебалансовой задолженности и 3) требование дополнительных залоговых гарантий?

## **4. Раскрытие информации**

32. Подлежат ли раскрытию в открытых источниках требования банка, относящиеся к CRM?

## **С. Валидация информационных технологий, задействованных в рейтинговой системе**

В соответствии с Международной конвергенцией измерения капитала и стандартов капитала (Базель 2) термин «рейтинговая система» включает в себя все методы, процедуры, системы контроля и сбора данных, а также информационные системы, которые содействуют оценке кредитного риска, присвоению внутренних рейтингов и количественным оценкам вероятности дефолта и убытка. Таким образом, валидация информационных технологий (ИТ), задействованных в рейтинговой системе должна являться неотъемлемой частью общего процесса валидации.

Данный раздел документа содержит рекомендации по оценке ИТ задействованных в процессе разработки и валидации модели, а также внедрении ее в промышленную среду и эксплуатации и является приложением к Методическим рекомендациям по реализации подхода к расчету кредитного риска на основе внутренних рейтингов банков.

ИТ валидация может быть разделена на две части:

- Анализ ИТ, связанный с разработкой модели
- Анализ ИТ, связанный с реализацией модели в промышленной среде и ее использованием

Оба исследования должны быть проведены в рамках первоначальной валидации. Исследования могут быть проведены различными подразделениями, обладающими достаточным уровнем компетенции для предоставления адекватной оценки.

Ответы на тестовые вопросы целесообразно структурировать следующим образом:

- Описание ситуации «как есть»: описание текущей ситуации в соответствии с имеющимися документами и опроса ответственных лиц. Описание должно быть исчерпывающим и понятным третьим лицам (ответы вида «да»/ «нет» недостаточны).

- Уровень адекватности: оценка адекватности в соответствии с «принципом светофора»: «зеленый» - полное выполнение всех нормативных требований, «желтый» - существуют некоторые отклонения/ недостатки, «красный» может означать наличие не только некоторых недостатков, но и некорректное использование рейтинговой системы. Тем не менее, «желтый» и «красный» результат выполнения теста может приводить к рекомендациям как высокого, так и среднего и низкого приоритета в зависимости от степени существенности выявленных недостатков. Необходимо определение критерия материальности, оценивающего в количественном выражении воздействие на 1) повышение стоимости риска 2) возможное повышение объема балансовой и внебалансовой задолженности, что приведет к увеличению PD и LGD

- Анализ недостатков: исчерпывающее и подробное описание возможных недостатков в зависимости от требований теста. Должны быть определены бизнес-кейсы по различным альтернативам (альтернативные издержки), например сокращение портфеля с последующим сокращением объема задолженности/лимитов.

- Рекомендация: предложение определенных действий, позволяющих исправить выявленные недостатки. Степень значимости (приоритет) рекомендации также должен быть указан.

- Список документов: перечень документов, задействованных в подготовке ответа на тестовый вопрос.

Для каждого из тестовых вопросов указано, с какой периодичностью необходимо проводить соответствующий анализ. По результатам валидации должен быть представлен соответствующий отчет. Выполнение рекомендаций должно быть закреплено в плане действий с четким указанием ответственных лиц и сроков исполнения. Статус выполнения рекомендаций также оценивается в рамках ежегодной валидации.

## 1. Валидация ИТ, связанных с моделью

### 1.1. Среда разработки

Вопрос		Первонач.	Периодич.
<i>ИТ архитектура</i>			
1	Опишите ИТ архитектуру, задействованную в процессе разработки модели подразделением, ответственным за разработку моделей? Вовлечено ли официально ИТ подразделение банка в ее управление?	X	
<i>Инфраструктура технологий ИТ и безопасность</i>			
2	Включена ли среда разработки в План обеспечения непрерывной деятельности банка и План восстановления сред?	X	
3	Какие применяются правила безопасности? Каким образом осуществляется управление процессом авторизации? Каким образом обеспечивается разграничение прав доступа и ролей?	X	
<i>Управление данными и моделью</i>			
4	Перечислите все источники данных, использованных в разработке. Каждая ли выборка данных сохранена единственным образом и обеспечено соответствующее хранение? Включите описание приложений, обеспечивающих резервное копирование, хранение и версию.	X	
5	Все ли программные коды разработки соответствующим образом организованы и документированы, включая промежуточные шаги, финальные результаты и результаты статистических тестов? Сохранены ли все версии программ? Включите описание приложений, обеспечивающих резервное копирование, хранение и версию.	X	
6	Возможно ли воспроизвести все шаги разработки, промежуточные и финальные результаты? Приведите доказательства выполненных тестов.	X	

Вопрос		Первонач.	Периодич.
<i>Управление ИТ знаниями</i>			
7	Планируется ли проведение приемочного тестирования, сопровождающееся подробным списком необходимых для выполнения тестов? Какие подразделения вовлечены в тестирование? Имеется ли согласованный план тестирования с четко определенными ролями и зонами ответственности?	X	
8	Запланировано ли в рамках приемочного тестирования проведение проверки качества работы пакетных процедур (например, устаревание рейтинга, события, связанные с изменениями в клиентском файле и т.д.)? Проверяется ли весь процесс присвоения рейтинга, включая проверку распределения ролей и корректности функционирования на каждом шаге?	X	
9	Адекватно ли описаны и документированы все тесты и результаты, полученные в рамках приемочного тестирования? Получено ли официальное одобрение от бизнеса на внедрение программного обеспечения в промышленную эксплуатацию? Обеспечена ли версияемость и сохранность документации? Есть ли в распоряжении банка стандартный инструментарий тестирования (напр., HP Quality Center) для систематического отслеживания и контроля хода тестирования и создания его документации для других целей (напр., требования аудита)?	X	X
10	Были ли подготовлены бизнес/функциональные спецификации на рейтинговую систему/модель? Были ли они предоставлены ответственному ИТ подразделению? Ести ли уверенность в том, что ИТ подразделение поняло требования? Обеспечена ли версияемость и сохранность документации?	X	
11	Содержит ли бизнес/функциональная спецификация информацию о процедурах, внедряемых в пакетном режиме (например, как и когда срабатывает ограничение рейтинга по возрасту, события, связанные с клиентской информацией, и т.п.)? Четко ли распределены роли и шаги процесса присвоения рейтинга?	X	

## 1.2. Валидация ИТ платформы

Вопрос		Первонач.	Периодич.
<i>ИТ процессы</i>			
1	Определено ли и закреплено ли подразделение, осуществляющее ответы на запросы в области ИТ?	X	X
2	Каким образом осуществляется информирование и коммуникация по ИТ запросам? Существует ли для этого специальная система?	X	X
3	Существует ли Орган, контролирующий и утверждающий приоритеты и ИТ работы, связанные с рейтинговыми системами? Закреплен ли документально порядок работы данного Органа?	X	X
4	Существует ли организованный процесс обновления статусов выполнения ИТ работ с четко определенными ролями и зонами ответственности? Предусмотрена ли прозрачная процедура эскалации на случай задержек или появлении рисков внедрения?	X	X

## 1.3. Промышленная среда

Вопрос		Первонач.	Периодич.
<i>ИТ архитектура</i>			
5	Опишите ИТ архитектуру процесса расчета рейтинга. Является ли ИТ подразделение официально ответственным за ее управление?	X	X
6	Какие ИТ приложения интегрированы с ИТ платформой? По каким правилам/интерфейсам прочие приложения взаимодействуют с ИТ платформой? Опишите основные потоки данных.	X	X
7	Какие контрольные процедуры внедрены для проверки целостности потоков данных между ИТ приложениями?	X	X
8	Поддерживает ли ИТ платформа процесс расчета рейтинга в соответствии с установленными ролями и шагами? Все ли шаги автоматизированы?	X	X

Вопрос		Первонач.	Периодич.
	Осуществляется ли автоматическое информирование участников процесса для поддержки взаимодействия между различными ролями?		
<i>Инфраструктура технологий ИТ и безопасность</i>			
9	Включена ли промышленная среда в План обеспечения непрерывной деятельности банка и План восстановления сред?	X	X
10	Какие применяются правила безопасности? Каким образом осуществляется управление процессом авторизации? Каким образом обеспечивается разграничение прав доступа и ролей? Обеспечивается ли журналирование действий пользователя?	X	X
<i>Управление данными и моделью</i>			
11	Перечислите все источники данных, используемых для расчета рейтинга. Все ли используемые базы данных сохранены единственным и надлежащим образом? Включите описание приложений, обеспечивающих резервное копирование, хранение и версию.	X	X
12	Все ли компоненты программного обеспечения хранятся соответствующим образом с обеспечением версии? Включите описание приложений, обеспечивающих резервное копирование, хранение и версию.	X	X
13	Возможно ли, используя сохраненные данные и соответствующую версию программного обеспечения, воспроизвести алгоритм расчета модели и результаты за предыдущую дату/период? Существует ли соответствующая процедура, обеспечивающая возможность воспроизведения расчета рейтинга? Приведите результаты выполненных тестов.	X	X
14	Каким образом организован процесс передачи модели от подразделения разработки ИТ подразделению для ее внедрения в промышленную среду? Какие инструменты используются для выполнения этих работ?	X	
<i>Управление ИТ знаниями</i>			
15	В случае получения компонентов программного обеспечения от внешних поставщиков, была ли предоставлена адекватная и подробная документация, инструкция пользователя и техническая спецификация? Все ли документы сохранены и обеспечена версия? Есть ли у банка возможность влиять на разработку программного обеспечения в случае привлечения внешнего поставщика?	X	X
16	Доступны ли все технические/архитектурные спецификации и описания словарей данных? Все ли они сохранены с указанием версий? Согласуются ли они с текущей версией программного обеспечения в промышленной среде и актуальными версиями бизнес/технических спецификаций?	X	X



# Приложения

Приложение 1

## 1. Пороговые значения для интерпретации результатов количественных тестов

В Приложении 1 приведены предлагаемые пороговые значения<sup>13</sup> для интерпретации результатов количественных тестов. При составлении банком своего справочника пороговых значений необходимо учитывать, что приведенный перечень нельзя считать исчерпывающим. В описании пороговых значений для тестов используются следующие понятия:

- Модель – метод оценки вероятности дефолта PD клиента/ транзакции для портфеля, который является отдельным классом активов (корпоративный, розничный, банки и т.п.). Возможна разработка модели для портфеля (части класса активов) при наличии обоснования и четкого определения критериев такого портфеля (например, крупные и средние корпоративные клиенты). Портфели в рамках одного класса активов могут формироваться в зависимости от размера клиентов, уровня дефолта, установленного рейтингового процесса, ИТ инфраструктуры и т.п. Модель, как правило, состоит из модулей.

- Модуль – компонент модели, состоящий из комбинации факторов, и отличающийся от других модулей своими качественными характеристиками (например, финансовый модуль, качественный модуль, поведенческий модуль и т.п.). Результат работы модуля может быть как PD для конкретного модуля, так и скоринговый балл.

- Суб-модуль (только для портфелей ссуд физическим лицам) – часть модуля, обычно заявочного, представляющая собой скоринговую карту, построенную, как правило, на основании запрашиваемого продукта.

- Сегмент – группа клиентов/ транзакций, для которых значимые факторы модели/ модуля отличаются полностью или частично. Например, для корпоративных клиентов сегментация может быть сделана на основании размера компании, индустрии; для розничных клиентов – продукта. Калибровка также может быть выполнена с учетом сегментации.

### *Интерпретация результатов тестов ROPM*

Тест	Результат		Уровень значимости			
	желтый	красный	высокий	средний	низкий	недостов.
<i>Высокодефолтный корпоративный портфель</i>						
Модель в целом на этапе разработки	<60%	<50%	<10%	<20%	<40%	>40%
Сегменты на этапе разработки	<45%	<35%	<10%	<20%	<40%	>40%
Модули на этапе разработки, кроме качественного	<45%	<35%	<10%	<20%	<40%	>40%
Качественный модуль на этапе разработки	<30%	<20%	<10%	<20%	<40%	>40%
Факторы на этапе разработки	<15%	<10%	<10%	<20%	<40%	>40%

<sup>13</sup> Приводимые пороговые значения являются ориентировочными. В любом случае интерпретация полученных результатов должна основываться на текущей экономической ситуации, состоянии портфеля, планах по развитию бизнеса, экспертном мнении, заложенном при построении модели.

Тест	Результат		Уровень значимости			
	желтый	красный	высокий	средний	низкий	недостов.
Модель в целом на этапе валидации	<55%	<45%	<10%	<20%	<40%	>40%
Сегменты на этапе валидации	<40%	<30%	<10%	<20%	<40%	>40%
Модули на этапе валидации, кроме качественного	<35%	<25%	<10%	<20%	<40%	>40%
Качественный модуль на этапе валидации	<25%	<15%	<10%	<20%	<40%	>40%
Факторы на этапе валидации	<10%	<5%	<10%	<20%	<40%	>40%
Сравнение модели на этапах разработки и валидации	>10%	>20%	<10%	<20%	<40%	>40%
Сравнение сегментов на этапах разработки и валидации	>10%	>20%	<10%	<20%	<40%	>40%
Сравнение модулей на этапах разработки и валидации	>10%	>20%	<10%	<20%	<40%	>40%
Сравнение факторов на этапах разработки и валидации (относительное)	>10%	>20%	<10%	<20%	<40%	>40%
<i>Высокодефолтный портфель ссуд физическим лицам</i>						
Модель в целом на этапе разработки	<65%	<55%	<10%	<20%	<40%	>40%
Сегменты на этапе разработки	<55%	<45%	<10%	<20%	<40%	>40%
Модули на этапе разработки:			<10%	<20%	<40%	>40%
С поведенческим компонентом	<60%	<50%				
Без поведенческого компонента	<45%	<35%				
Суб-модули на этапе разработки:			<10%	<20%	<40%	>40%
С поведенческим компонентом	<60%	<50%				
Без поведенческого компонента	<35%	<25%				
Поведенческие факторы на этапе разработки	<20%	<15%	<10%	<20%	<40%	>40%
Прочие факторы на этапе разработки	<15%	<10%	<10%	<20%	<40%	>40%
Модель в целом на этапе валидации	<60%	<50%	<10%	<20%	<40%	>40%
Сегменты на этапе валидации	<50%	<40%	<10%	<20%	<40%	>40%

Тест	Результат		Уровень значимости			
	желтый	красный	высокий	средний	низкий	недостов.
Модули на этапе валидации: С поведенческим компонентом	<55%	<45%	<10%	<20%	<40%	>40%
Без поведенческого компонента	<40%	<30%				
Суб-модули на этапе валидации: С поведенческим компонентом	<55%	<45%	<10%	<20%	<40%	>40%
Без поведенческого компонента	<30%	<20%				
Поведенческие факторы на этапе валидации	<15%	<10%	<10%	<20%	<40%	>40%
Прочие факторы на этапе валидации	<10%	<5%	<10%	<20%	<40%	>40%
Сравнение модели на этапах разработки и валидации	>10%	>20%	<10%	<20%	<40%	>40%
Сравнение сегментов на этапах разработки и валидации	>10%	>20%	<10%	<20%	<40%	>40%
Сравнение модулей на этапах разработки и валидации	>10%	>20%	<10%	<20%	<40%	>40%
Сравнение суб-модулей на этапах разработки и валидации	>10%	>20%	<10%	<20%	<40%	>40%
Сравнение факторов на этапах разработки и валидации (относительное)	>10%	>20%	<10%	<20%	<40%	>40%

**Интерпретация результатов тестов на концентрацию  
(индекс Херфиндаля)**

Тест	Результат	
	желтый	красный
Распределение текущего портфеля по рейтинговой шкале	>20%	>30%
Распределение текущего портфеля на уровне модуля по рейтинговой шкале	>20%	>30%
Распределение сегментов текущего портфеля по рейтинговой шкале	>20%	>30%
Изменение распределения портфеля (сегмента) по рейтинговой шкале с этапа разработки до этапа валидации	>20%	>30%

**Интерпретация результатов тестов на репрезентативность/ стабильность  
(индекс PSI)**

Тест	Результат		Уровень значимости		
	желтый	красный	высокий	средний	низкий
<i>Репрезентативность</i>					
Сравнение выборки для разработки (в т.ч. для калибровки) с текущим портфелем	>0.1	>0.2	<3%	<10%	<30%
Сравнение на уровне модулей в разрезе сегментов выборки для разработки и текущего портфеля	>0.1	>0.2	<3%	<10%	<30%
Сравнение выборки для валидации и текущего портфеля в разрезе сегментов	>0.1	>0.2	<3%	<10%	<30%
<i>Стабильность</i>					
Сравнение распределения выборки для разработки (в т.ч. выборки для калибровки) и выборки для валидации в разрезе сегментов по рейтинговой шкале	>0.1	>0.2	<3%	<10%	<30%
Сравнение распределения выборки для разработки с выборкой для валидации на уровне модулей по рейтинговой шкале	>0.1	>0.2	<3%	<10%	<30%

## 2. Показатели качественной валидации

### 2.1. AUROC – площадь под ROC кривой

ROC кривая – это кривая, которая определяется следующим образом: предположим, что все «плохие»<sup>14</sup> и «хорошие»<sup>15</sup> наблюдения упорядочены в соответствии с присвоенным скоринговым баллом от худшего к лучшему. Тогда точки ROC кривой определяются как соответствие доли «хороших» наблюдений доле «плохих» наблюдений, в которой присвоенные скоринговые баллы хуже, чем лучший скоринговый балл из первой группы. Например, если лучший скоринговый балл в 10% «хороших» наблюдений с наихудшими скоринговыми баллами составляет 123 и доля «плохих» наблюдений со скоринговыми баллами хуже 123 составляет 20%, то точка ROC будет определена как соответствие 10% по оси долей «хороших» наблюдений и 20% по оси долей «плохих наблюдений».

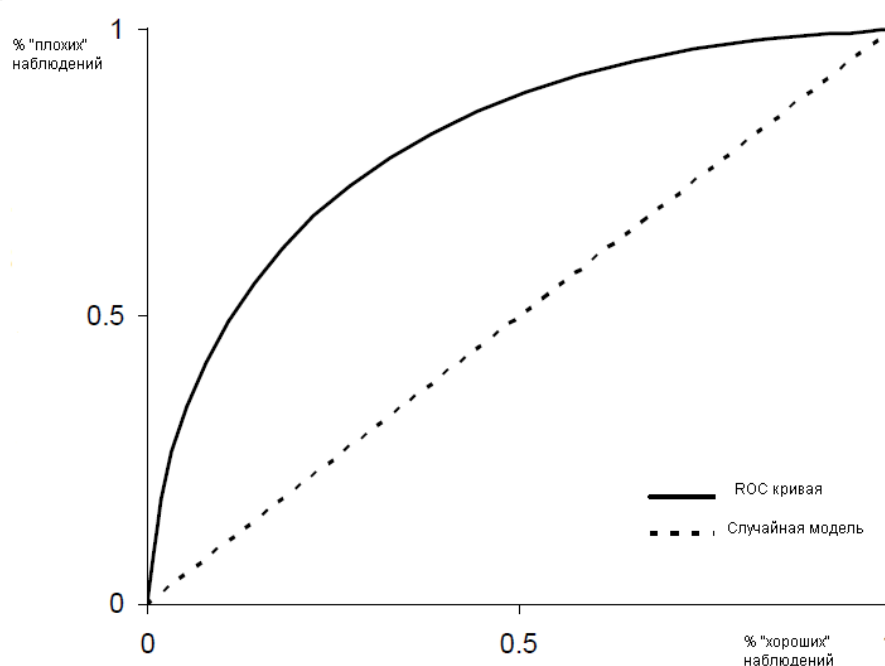


Рис.1. ROC кривая

Площадь под ROC кривой является мерой качества ранжирования. Она варьируется от 0 (модель работает с точностью наоборот – все «плохие» наблюдения получают лучший скоринговый балл, чем «хорошие») до 1 (идеальная модель – все «плохие» наблюдения получают худший скоринговый балл, чем «хорошие»). Если модель не может отделить «плохие» наблюдения от «хороший», значение AUROC будет равно 0.5.

<sup>14</sup> «Плохое» наблюдение –наблюдение испытавшее дефолт (или другое иное событие, отрицательно характеризующее это событие, например, достижение определенного количества дней просрочки) в заданном периоде (например, 12 месяцев).

<sup>15</sup> «Хорошее» наблюдение – наблюдение, не являющееся «плохим».

Простая формула для вычисления AUROC базируется на правиле трапеции для интегрирования по ROC кривой:

$$\hat{A} = \frac{\sum_{i,j} B_i \cdot G_j \cdot \alpha_{ij}}{B \cdot G}$$

где  $\hat{A}$  – несмещенная оценка площади под ROC кривой

$$\alpha_{ij} = \begin{cases} 1, & \text{если } i < j \\ 0.5, & \text{если } i = j \\ 0, & \text{если } i > j \end{cases}$$

$B_i, G_j$  – количество «плохих» и «хороших» наблюдений для  $i$ -го и  $j$ -го аргумента соответственно.

$B, G$  – общее количество «плохих» и «хороших» наблюдений соответственно.

Эта формула остается также актуальной и для непрерывной шкалы скоринговых баллов.

Приведенная оценка AUROC является несмещенной и ниже приведена формула для расчета соответствующей стандартной ошибки:

$$\hat{\sigma}_{AUROC} = \sqrt{\frac{\left(1 - \frac{\sum_i B_i G_i}{GB}\right) + \frac{(B-1)}{B^2 G} \sum_{i,j,k} B_i B_j G_k \beta_{ijk} + \frac{(G-1)}{G^2 B} \sum_{i,j,k} G_i G_j B_k \beta_{ijk} - 4(B+G-1)(\hat{A}-0.5)^2}{4(B-1)(G-1)}}$$

где  $\hat{A}$  – оценка площади под ROC кривой (AUROC).

$$\beta_{ijk} = \begin{cases} 1, & \text{если } k > i, j \text{ или } k < i, j \\ -1, & \text{если } i < k < j \text{ или } j < k < i \\ 0 & \text{в противных случаях} \end{cases}$$

$B_i, G_j$  – количество «плохих» и «хороших» наблюдений для  $i$ -го и  $j$ -го аргумента соответственно.

$B, G$  – общее количество «плохих» и «хороших» наблюдений соответственно.

Если шкала скоринговых баллов непрерывна, можно использовать приближительную формулу вычисления стандартной ошибки:

$$\hat{\sigma}_{AUROC} = \sqrt{\frac{\hat{A}(1-\hat{A}) + (B-1)\left(\frac{\hat{A}}{2-\hat{A}} - \hat{A}^2\right) + (G-1)\left(\frac{2\hat{A}^2}{1+\hat{A}} - \hat{A}^2\right)}{BG}}$$

где  $\hat{A}$  – оценка площади под ROC кривой (AUROC).

$B, G$  – общее количество «плохих» и «хороших» наблюдений соответственно.

Доверительный интервал может быть определен следующим образом:

$$P\left(\hat{A} - \hat{\sigma}_{AUROC} \Phi^{-1}\left(1 + \frac{\alpha}{2}\right) < AUROC < \hat{A} + \hat{\sigma}_{AUROC} \Phi^{-1}\left(1 + \frac{\alpha}{2}\right)\right) = \alpha$$

где  $\hat{A}$  – оценка площади под ROC кривой (AUROC).

$\alpha$  – уровень значимости.

$\Phi^{-1}$  – обратная кумулятивная функция нормального распределения.

Как преимущество AUROC можно назвать получаемую несмещенную оценку, возможность определить доверительный интервал и независимость от количества «хороших» наблюдений. Недостатком использования этой меры является необходимость обязательной сортировки наблюдений перед расчетом.

## 2.2. Мера AR (Accuracy Ratio)

Мера AR тесно связана с AUROC и также является мерой качества ранжирования. Другое ее название – **индекс Джини** (GINI index, GINI). AR вычисляется как отношение двух площадей, определяемых с помощью кривой CAP (cumulative accuracy profile), строящейся по принципу, аналогичному ROC кривой. Пусть все наблюдения отсортированы по присвоенному скоринговому баллу от наихудшего к наилучшему. Тогда точки CAP кривой определяются соотношением доли наблюдений к такой доле «плохих» наблюдений, где скоринговый балл хуже, чем лучший балл любого наблюдения, включенного в последнюю группу наблюдений. То есть, если лучший скоринговый балл в 10% наблюдений с наихудшими скоринговыми баллами составляет 123 единицы и доля «плохих» наблюдений со скоринговым баллом хуже 123 составляет 20%, то точка кривой CAP будет определяться 10% по оси абсцисс и 20% по оси ординат.

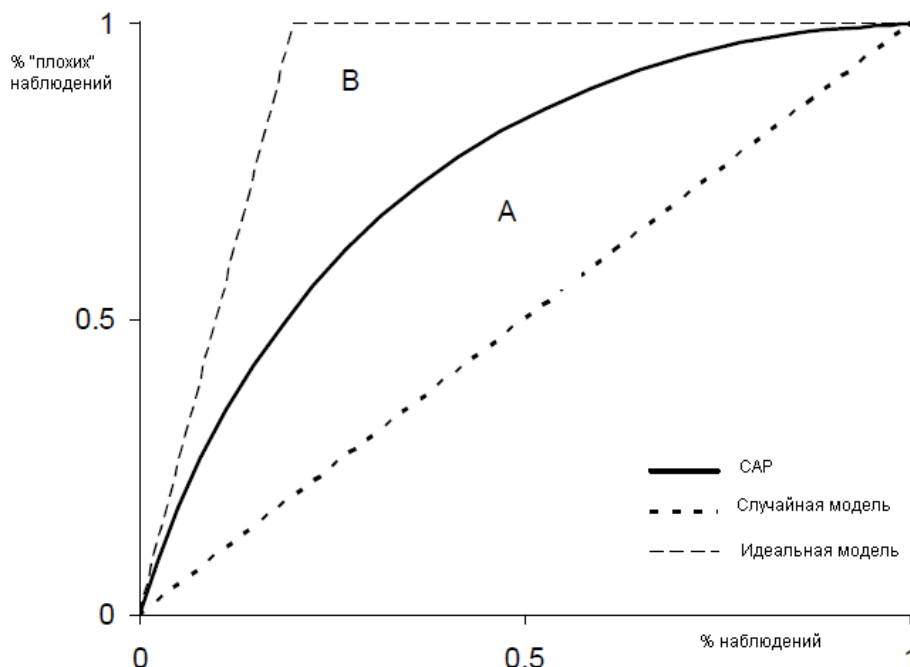


Рис.2. Мера Accuracy ratio

AR определяется как отношение площади А к площади В (см. Рисунок 2). Площадь А – площадь области между CAP кривой и CAP кривой, соответствующей случайной модели. Площадь В – площадь области между CAP кривой идеальной модели и CAP кривой случайной модели.

AR является хорошей мерой оценки качества ранжирования. Диапазон значений AR находится в границах от -1 (модель работает с точностью наоборот – все «плохие» наблюдения получают скоринговый балл лучше, чем любое из «хороших») до 1 (идеальная модель – все «плохие» наблюдения получают скоринговый балл ниже «хороших»). Если модель не отличает «хорошие» наблюдения от «плохих», AR=0.

Взаимосвязь AR и AUROC:

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

Также значение AR можно вычислить по простой формуле:

$$\hat{C} = \frac{\sum_{i,j} B_i G_j \gamma_{ij}}{BG}$$

где  $\hat{C}$  - несмещенная оценка AR.

$$\gamma_{ij} = \begin{cases} 1, & \text{если } i < j \\ 0, & \text{если } i = j \\ -1, & \text{если } i > j \end{cases}$$

$B_i, G_j$  – количество «плохих» и «хороших» наблюдений для  $i$ -го и  $j$ -го аргумента соответственно.

$B, G$  – общее количество «плохих» и «хороших» наблюдений соответственно.

Используя формулу связи AR и AUROC, можно получить стандартную ошибку для AR:

$$\hat{\sigma}_{AR} = \sqrt{\frac{\left(1 - \frac{\sum B_i G_i}{GB}\right) + \frac{(B-1)}{B^2 G} \sum_{i,j,k} B_i B_j G_k \beta_{ijk} + \frac{(G-1)}{G^2 B} \sum_{i,j,k} G_i G_j B_k \beta_{ijk} - (B+G-1)\hat{C}^2}{(B-1)(G-1)}}$$

где  $\hat{C}$  – оценка AR.

$$\beta_{ijk} = \begin{cases} 1, & \text{если } k > i, j \text{ или } k < i, j \\ -1, & \text{если } i < k < j \text{ или } j < k < i \\ 0 & \text{в противных случаях} \end{cases}$$

$B_i, G_j$  – количество «плохих» и «хороших» наблюдений для  $i$ -го и  $j$ -го аргумента соответственно.

$B, G$  – общее количество «плохих» и «хороших» наблюдений соответственно.

Также как для AUROC существует упрощенная формула расчета стандартной ошибки:

$$\hat{\sigma}_{AR} = \sqrt{\frac{1 - \hat{C}^2 + (B-1) \left(4 \frac{\hat{C}+1}{3-\hat{C}} - (\hat{C}+1)^2\right) + (G-1) \left(4 \frac{(\hat{C}+1)^2}{3+\hat{C}} - (\hat{C}+1)^2\right)}{BG}}$$

где  $\hat{C}$  – оценка AR.

$B, G$  – общее количество «плохих» и «хороших» наблюдений соответственно.

Доверительный интервал для AR вычисляется также как и для AUROC.

Достоинства и недостатки мер AR и AUROC одинаковы.

### 2.3. Оценка стандартной ошибки для индекса Джини

Помимо значения индекса Джини оценивается стандартная ошибка. Ее расчет основан на оценке стандартной ошибки площади под кривой ROC, известной также как статистика Манна-Уитни. Оценка дисперсии для площади под кривой ROC:

$$\hat{\sigma}_U^2 = \frac{1}{4(N_D - 1)(N_{ND} - 1)} \times$$



$$\times \left[ \mathbf{1} + (N_{ND} - 1)\widehat{P}_{D,D,ND} + (N_{ND} - 1)\widehat{P}_{ND,ND,D} - 4(N_D + N_{ND} - 1) \left( \widehat{U} - \frac{1}{2} \right)^2 \right]$$

где

$\widehat{U}$  – площадь под кривой ROC (статистика Манна-Уитни),

$N_D$  – количество дефолтов,

$N_{ND}$  – количество недефолтов,

$\widehat{P}_{D,D,ND}$  и  $\widehat{P}_{ND,ND,D}$  – оценки величин  $P_{D,D,ND}$  и  $P_{ND,ND,D}$ :

$$\begin{aligned} P_{D,D,ND} &= P(S_{D,1}, S_{D,2} < S_{ND}) + P(S_{ND} < S_{D,1}, S_{D,2}) \\ &\quad - P(S_{D,1} < S_{ND} < S_{D,2}) - P(S_{D,2} < S_{ND} < S_{D,1}) \\ P_{ND,ND,D} &= P(S_{ND,1}, S_{ND,2} < S_D) + P(S_D < S_{ND,1}, S_{ND,2}) \\ &\quad - P(S_{ND,1} < S_D < S_{ND,2}) - P(S_{ND,2} < S_D < S_{ND,1}) \end{aligned}$$

где  $P(\omega)$  – вероятность события  $\omega$ ,

$S_{D,1}, S_{D,2}$  – скоринговые баллы по двум наблюдениям из выборки дефолтных заемщиков, выбранным случайно и независимо друг от друга,

$S_{ND,1}, S_{ND,2}$  – скоринговые баллы по двум наблюдениям из выборки недефолтных заемщиков, выбранным случайно и независимо друг от друга,

$S_{ND}$  – скоринговый балл по наблюдению из выборки недефолтных заемщиков, выбранному случайно и независимо от выбора наблюдений с баллами  $S_{D,1}$  и  $S_{D,2}$ ,

$S_D$  – скоринговый балл по наблюдению из выборки дефолтных заемщиков, выбранному случайно и независимо от выбора наблюдений с баллами  $S_{ND,1}$  и  $S_{ND,2}$

Индекс Джини (ассигасу ratio) равен

$$AR = 2(\widehat{U} - 0.5)$$

Оценка дисперсии индекса и стандартной ошибки Джини:

$$\widehat{\sigma}_{AR}^2 = 4\widehat{\sigma}_{\widehat{U}}^2, \quad \widehat{\sigma}_{AR} = 2\widehat{\sigma}_{\widehat{U}}.$$

Малое значение стандартной ошибки индекса Джини означает, что данных достаточно для достаточно точной оценки индекса Джини, т.е. вероятность значительного отклонения истинного индекса Джини от оцененного по выборке мала. Если же в силу недостатка данных стандартная ошибка достаточно велика (свыше 5%), с пороговыми значениями сравнивается верхняя граница доверительного интервала для индекса Джини, который строится по формуле

$$CI_{AR} = \left\{ AR: \widehat{AR} + \widehat{\sigma}_{AR}\Phi^{-1}\left(\frac{\alpha}{2}\right) < AR < \widehat{AR} + \widehat{\sigma}_{AR}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \right\}$$

где  $AR$  – значения индекса Джини, принадлежащие интервалу,

$\widehat{AR}$  – оценка AR по выборке,

$\alpha$  – уровень значимости,

$\Phi^{-1}$  – обратная кумулятивная функция нормального распределения.

Уровень значимости принимается равным 0.05, поскольку большие значения означают высокую вероятность присвоить красный цвет модели, которая в действительности обеспечивает качество ранжирования не ниже порогового значения. Иными словами, при сравнении выборочного индекса Джини с пороговым значением (в случае, когда стандартная

ошибка выше 5%) мы проверяем нулевую гипотезу о том, что истинное значение индекса не ниже порогового значения, принимая уровень значимости равным 0.05.

## 2.4. Критерий Колмогорова-Смирнова

Критерий согласия Колмогорова или Критерий согласия Колмогорова-Смирнова — статистический критерий, использующийся для определения того, подчиняются ли два эмпирических распределения одному закону, либо того, подчиняется ли полученное распределение предполагаемой модели.

Критерий Колмогорова-Смирнова о проверке гипотезы об однородности двух эмпирических законов распределения является одним из основных и наиболее широко используемых непараметрических критериев, так как достаточно чувствителен к различиям в исследуемых выборках.

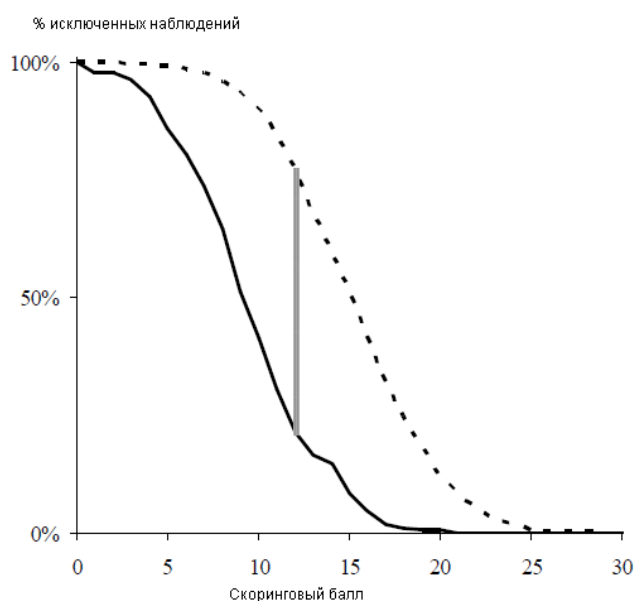


Рис.3. Критерий Колмогорова-Смирнова

На Рисунке 3 пунктирная (сплошная) линия получена путем вычитания из единицы эмпирического кумулятивного распределения «хороших» («плохих») наблюдений. Вертикальное расстояние между двумя распределениями и есть значение критерия Колмогорова-Смирнова. Однако, оценка является смещенной, что является недостатком этого критерия.

## 2.5. Мера Information value (IV)

Мера Information value (IV) также является мерой оценки качества ранжирования и вычисляется по следующей формуле:

$$IV = \sum_i IV_i = \sum_i (B\%_i - G\%_i) \cdot \log \left( \frac{B\%_i}{G\%_i} \right)$$

где  $IV_i$  –  $i$ -ое значение IV.

$B\%_i = \frac{B_i}{\sum_j B_j}$  где  $B_i$  – количество «плохих» наблюдений для  $i$ -го атрибута.

$G\%_i = \frac{G_i}{\sum_j G_j}$  где  $G_i$  – количество «хороших» наблюдений для  $i$ -го атрибута.

IV может принимать значения от 0 до бесконечности, при этом низкие значения соответствуют низкой различительной способности модели.

Достоинством этой меры является простота расчета и возможность оценки различительной способности для каждого фактора, но при этом оценка является смещенной.

Стоит отметить, что IV также называют PSI (population stability index – индекс стабильности популяции), который используется для определения степени различия между двумя выборками (например, для проверки репрезентативности).

### 3. Инструменты оценки стабильности моделей LGD/EAD

В целях оценки стабильности моделей для LGD и EAD применяются одни и те же методы, описанные ниже.

#### 3.1. Индекс стабильности системы

Назначение индекса стабильности системы – определить, одинаково ли распределены две дискретные выборки. Преимущество данного теста в том, что не делается никаких предположений о конкретном виде распределения. При этом недостатком является то, что сравниваться могут только дискретные выборки. Если выборка распределена непрерывно, то для применения данного метода необходимо провести разбиение данных на некоторые интервалы, границы которых должны быть определены экспертным путем.

Значение индекса определяется следующей формулой:

$$SSI = \sum_i (B_i - A_i) \times \ln\left(\frac{B_i}{A_i}\right)$$

где  $A_i$  и  $B_i$  представляют собой долю элементов из выборок А (валидируемые значения) и В («эталон»), которые принадлежат интервалу  $i$ .

Критерий вынесения суждения об одинаковости распределения:

$SSI < 0.1$	Нет различий
$SSI \in [0.1; 0.2]$	Небольшие различия
$SSI > 0.2$	Существенные различия

##### 3.1.1. Критерий согласия Колмогорова-Смирнова

Критерий согласия Колмогорова-Смирнова используется для того, чтобы определить, относятся ли две выборки к одной и той же генеральной совокупности.

Формулируются следующие гипотезы:

- H<sub>0</sub>: Распределения одинаковы
- H<sub>1</sub>: Распределения разные

Предположений о конкретном виде распределения не делается. Критерий основан на вычислении максимального расстояния между двумя кумулятивными функциями распределения.

Критерий основывается на трех предположениях:

- $X_i$  и  $Y_i$  - независимые случайные выборки, которые имеют кумулятивные распределения  $F_X$  и  $F_Y$
- $F_X$  и  $F_Y$  должны быть непрерывными
- Возможно упорядочивание данных

Для расчета максимального расстояния строятся два кумулятивных распределения:

$$F_X(x) = \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\{X_i < x\}} \quad F_Y(x) = \frac{1}{M} \sum_{i=1}^M \mathbf{1}_{\{Y_i < x\}}$$

Где  $X_i$ - элемент из выборки X ( $i=1, \dots, N$ ),  $Y_i$ - элемент из выборки Y ( $i=1, \dots, M$ ). Максимальное расстояние определяется как

$$D_{MN} = \max_x |F_X(x) - F_Y(x)|$$

Тестовая статистика определяется следующим образом:

$$T = \sqrt{\frac{NM}{N+M}} \times D_{MN}$$

Когда объем выборок стремится к бесконечности, величина  $T$  распределена по закону Колмагорова-Смирнова. Гипотеза  $H_0$  отвергается, если  $T > K_\alpha$ , где  $K_\alpha$  - квантиль распределения Колмогорова-Смирнова для уровня значимости  $\alpha$ .

Размер выборки считается достаточным для использования данного метода, если в ней более 40 элементов.

## 3.2. Инструменты оценки дискриминационной способности моделей LGD

Оценка дискриминационной способности моделей LGD может производиться на двух уровнях:

- Уровень модели (используются все данные, доступные для валидации)
- Уровень групп (данные для валидации группируются определенным образом, после чего оценка дискриминационной способности строится на основе групповых значений)

### 3.2.1. Loss Capture Ratio – Коэффициент покрытия убытков<sup>16</sup>

Целью данного метода является оценка способности модели правильно упорядочивать значения LGD, где критерием является покрытие совокупного убытка в следствие реализации кредитного риска. Валидируемая модель соотносится с идеальной и случайной моделями.

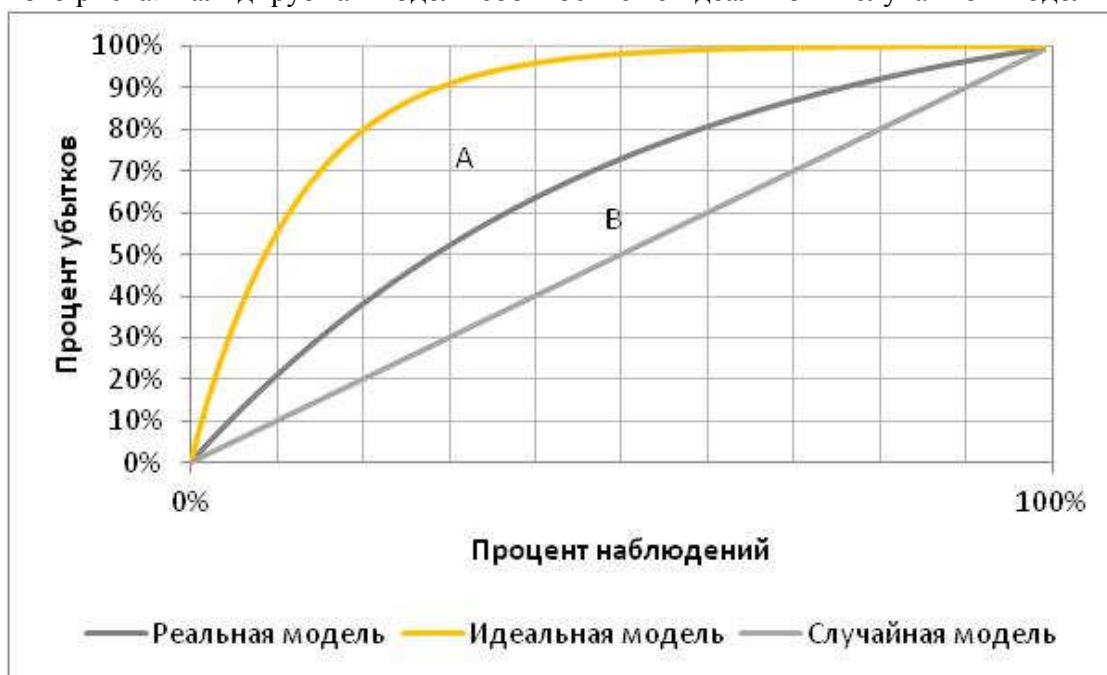


Рис.5. Loss Capture Curve

Для применения метода строится кривая покрытия убытков («Loss Capture Curve») аналогичная по своему смыслу ROC-кривой (рис.Рис.5).

Для построения кривой покрытия убытков ряд предсказанных значений LGD упорядочивается по убыванию. Затем для каждого предсказанного значения LGD рассчитывается кумулятивная функция доли реально зафиксированного убытка:

$$F_0 = 0 \qquad F_i = F_{i-1} + \frac{EAD_i \times LGD_i^{Real}}{\sum_i EAD_i \times LGD_i^{Real}}$$

<sup>16</sup> Validation techniques and performance metrics for loss given default models, The Journal of Risk Model Validation (3–26)

где  $EAD_i \times LGD_i^{Real}$  - реальный убыток по i-ому наблюдению.

Под идеальной моделью предполагается модель, которая безошибочно упорядочивает наблюдаемые значения LGD. Кривая для идеальной модели строится аналогично реальной кривой, за исключением того, что упорядочивается ряд не прогнозных, а реализовавшихся значений LGD.

Под случайной моделью подразумевается гипотетическая модель, которая сортирует LGD случайным образом. Кривая для случайной модели представляет собой прямую, выходящую из точки (0;0) под углом 45 градусов.

Коэффициент покрытия убытков находится как отношение площадей фигур B и A+B. Чем ближе коэффициент к 1, тем лучше дискриминационная способность модели.

### 3.2.2. Cumulative LGD Accuracy Ratio (CLAR) – кумулятивный коэффициент точности LGD

Данный коэффициент является мерой способности модели правильно упорядочивать значения LGD. Подход CLAR применяется на уровне групп и требует разбивки значений LGD на интервалы. Для каждого интервала k вычисляется значение кумулятивной функции:

$$P_k = \sum_{i=1}^k \frac{N_{i\text{correct}}}{N}$$

где  $N_{i\text{correct}}$  – число элементов, для которых и предсказанное и реализованное значения LGD попали в i-тый интервал.

N – общее число наблюдений

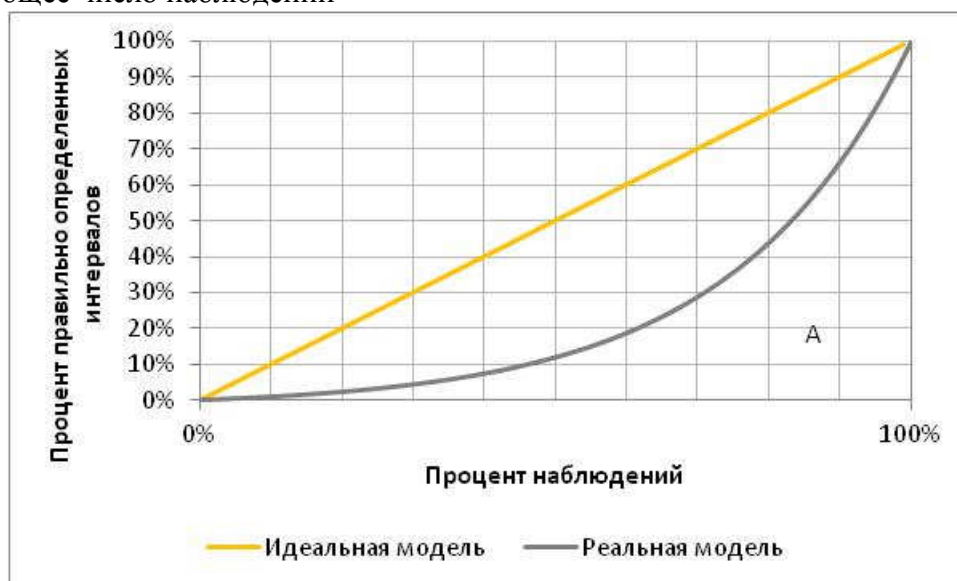


Рис.6. Кривая CLAR

Площадь под графиком кумулятивной функции (A) определяет дискриминационную способность модели (рис.Рис.6):  $CLAR = 2A$

Идеальной является модель, для которой все реализовавшиеся значения LGD оказались в тех же группах, что и предсказанные. Кривая для идеальной модели представляет собой прямую, выходящую из точки (0;0) под углом 45 град. CLAR для идеальной модели равен 1.

### 3.2.3. Коэффициент ранговой корреляции Спирмена

Коэффициент корреляции Спирмена является непараметрическим тестом, который проверяет наличие корреляции между двумя выборками. Формулируются гипотезы:

H0: Не существует полож. корреляции между предсказанными значениями LGD и наблюдаемыми

H1: Существует полож. корреляция между предсказ. значениями LGD и наблюдаемыми

Тестовая статистика основана на ранкинге двух выборок: выборке наблюдаемых значений LGD и выборке предсказанных значений. Корреляция определяется:

$$r_s = 1 - \frac{6}{N(N^2 - 1)} \sum_{i=1}^N (u_i - v_i)^2$$

где  $u_i$  - ранг значения LGD, предсказанного моделью

$v_i$  - ранг соответствующего наблюдаемого LGD

$N$  - число наблюдений

Если  $r_s = 1$ , то модель оптимальна и ранкинги одинаковы

Если  $r_s = 0$ , то модель ранжирует выборку случайным образом

Если  $r_s = -1$ , то модель ранжирует LGD в обратном порядке

Если число наблюдений превышает 40, то распределение величины  $r_s \sqrt{N-1}$  стремится к нормальному. Таким образом, нулевая гипотеза может быть отклонена на уровне значимости  $\alpha$ , если

$$r_s \sqrt{N-1} > z_\alpha$$

где  $z_\alpha$  - квантиль нормального распределения.

Соблюдение данного неравенства означает статистически значимую корреляцию. В то же время, следует также анализировать само значение  $r_s$ .

### 3.3. Инструменты оценки качества калибровки моделей LGD/EAD

#### 3.3.1. Loss Shortfall LGD<sup>17</sup>

Оценка качества калибровки модели может быть проведена как путем сравнения реализовавшихся убытков (Loss) и убытков, предсказанных моделью (LGD×EAD), так и сравнением прогнозного и реального значения LGD. Коэффициент Loss Shortfall показывает, насколько потери от дефолта ниже, чем было предсказано моделью.

$$\text{Loss Shortfall} = 1 - \frac{\sum_{i=1}^N (\text{LGD}_i \times \text{EAD}_i)}{\sum_{i=1}^N (\text{OLGD}_i \times \text{EAD}_i)}$$

где  $\text{OLGD}$  = Observed LGD – реализовавшееся значение LGD

LGD – значение LGD, предсказанное моделью.

При значениях  $\text{Loss Shortfall} > 0$  или  $\text{Loss Shortfall} < -0.20$  модель в красной зоне.

При  $-0.20 \leq \text{Loss Shortfall} \leq -0.10$ , модель попадает в желтую зону

При  $-0.10 < \text{Loss Shortfall} \leq 0$ , модель попадает в зеленую зону

Данный показатель производит консервативную оценку, так как модель принимается только в тех случаях, когда наблюдаемые потери ниже, чем предсказанные.

---

<sup>17</sup> Backtesting Framework for PD, EAD and LGD, Master Thesis, Bauke Maarse

### 3.3.2. Стандартное абсолютное отклонение LGD

Данный метод анализирует абсолютную разницу между наблюдаемыми и предсказанными потерями и рассчитывается следующим образом:

$$MAD = \frac{\sum_{i=1}^N |OLGD_i - LGD_i| \times EAD_i}{\sum_{i=1}^N EAD_i}$$

При значениях  $MAD > 20\%$ , модель попадает в красную зону

При  $10\% \leq MAD \leq 20\%$ , модель попадает в желтую зону

При  $MAD < 10\%$ , модель попадает в зеленую зону

При применении показателей LossShortfall и MAD следует учитывать особенности распределения LGD. Распределение фактического LGD является бимодальным с максимумами вблизи нуля и единицы. Прогнозные же значения LGD могут быть расположены ближе к среднему значению в связи с требованиями несмещенности прогноза, что может приводить к значительным абсолютным отклонениям в случае более ярко выраженной бимодальности в фактических данных. Таким образом, в случае высоких значений MAD и/или Loss Shortfall, следует проанализировать, является ли этот факт следствием низкого качества модели, или возникает за счет бимодальности распределения фактических данных. Применение данных показателей рекомендуется в качестве вспомогательных величин.

### 3.3.3. Калибровка LGD/EAD при помощи критерия Стьюдента

Критерий используется для анализа степени отклонения наблюдаемых величин от предсказанных. Критерий применяется для анализа качества калибровки как LGD, так и EAD.

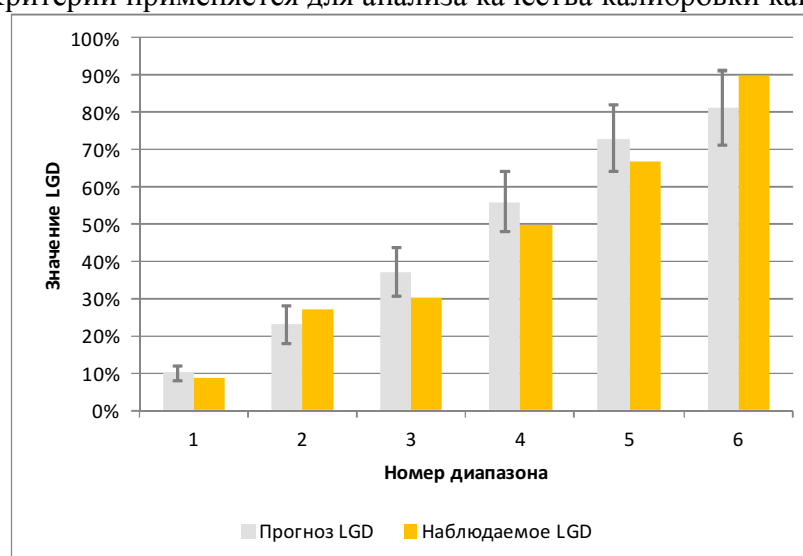


Рис.7. Калибровка LGD

Для применения данного критерия спрогнозированные значения исследуемой величины (например, LGD) сортируются по возрастанию и разбиваются на диапазоны. Затем для каждого диапазона вычисляется агрегированное значение прогнозного LGD и, с помощью критерия Стьюдента, строится доверительный интервал, в который при заданном уровне значимости, может попасть наблюдаемое значение (рис.Рис.7).

Диапазоны, для которых наблюдаемые значения выходят за рамки доверительных интервалов подлежат дополнительному анализу.



#### 4. Пороговые значения для интерпретации результатов количественных тестов

В таблице приведены предлагаемые пороговые значения для интерпретации результатов количественных тестов. При проведении валидации данные значения должны быть модифицированы банком с учетом специфики модели, характеристик портфеля, а также прочих факторов.

Пороговые значения для количественной оценки результатов бенчмаркинга банкам предлагается выбирать самостоятельно, основываясь на характеристиках модели и параметрах бенчмарка.

Тест	Желтая зона	Красная зона
Стабильность		
Индекс стабильности системы	>0.1	>0.2
Критерий Колмогорова-Смирнова	Принятие гипотезы на уровне 95%	Принятие гипотезы на уровне 85%
Дискриминационная способность LGD		
LCR	<50%	<25%
CLAR	<50%	<25%
Корреляция Спирмена	Принятие гипотезы на уровне 95%	Принятие гипотезы на уровне 90%
Качество калибровки		
Loss Shortfall*	От -20% до -10%	<-20% или >0
MAD*	>10%	>20%
T-критерий	Принятие гипотезы на уровне 95%	Принятие гипотезы на уровне 85%