University of Tartu

Faculty of Science and Technology

Institute of Mathematics and Statistics

Kristo Visk

Measures to assess the discriminatory power of loss given default models

Master's thesis (30 EAP)
Actuarial and Financial Engineering

Supervisor: Meelis Käärik, PhD

Tartu 2022

Mõõdikud hindamaks maksejõuetusest tingitud kahju prognoosivate mudelite järjestusvõimet

Magistritöö
Kristo Visk

Lühikokkuvõte. Magistritöö eesmärk on uurida mõõdikuid, mille abil saab hinnata maksejõuetusest tingitud kahju prognoosivate mudelite järjestusvõimet. Maksejõuetusest tingitud kahju prognoosivaid mudeleid kasutatakse finantsinstitutioonide mitteoodatava kahju kvantifitseerimisel. Esimeses peatükis antakse ülevaade Baseli Akordidest, regulatiivse kapitali hindamise nõuetest, mitteoodatava kahju hindamisest ning sisemiste riskihinnangute valideerimisest. Teises peatükis tutvustatakse erinevaid mõõdikud, mille abil saab hinnata maksejõuetusest tingitud kahju prognoosivate mudelite järjestusvõimet. Töös käsitletud mõõdikute hulka kuulub ka Euroopa Keskpanga sisemiste mudelite valideerimistulemuste raporteerimise juhendis välja toodud mõõde: üldistatud AUC. Töös analüüsitakse ja võrreldakse esitletud mõõdikute matemaatilisi omadusi. Lisaks pakutakse välja kaks täiendavat mõõdikut, mille abil on võimalik valideerimisprotsessi järeldusi toetada. Kolmandas peatükis viiakse läbi simulatsioon, mis illustreerib olukorda, kus enamus valideerimisvalimis olevaid vaatlusi kannab relatiivselt väikest kahju. Lisaks võrreldakse simulatsiooni tulemusi teises peatükis defineeritud mõõdikute alusel.

CERCS teaduseriala: P160 Statistika, operatsioonianalüüs, programmeerimine, finants- ja kindlustusmatemaatika.

Märksõnad. Krediidirisk, Baseli Akordid, sisereitingutel põhinevad mudelid, maksejõuetusest tingitud kahju prognoosivad mudelid, järjestusvõime, astakkorrelatsioon, üldistatud AUC.

Measures to assess the discriminatory power of loss given default models

Master's thesis
Kristo Visk

Abstract. The purpose of this master's thesis is to explore measures that can be used to assess the discriminatory power of loss given default models, which are used for the quantification of unexpected losses by financial institutions. In the first chapter, a general overview of the Basel Accords, the requirements related to regulatory capital calculations, the estimation of unexpected losses and the validation of internal risk estimates are provided. The second chapter highlights various measures that can be used to assess the discriminatory power of loss given default models, including a measure defined in the European Central Bank's instructions for reporting the validation results of internal models, which is referred to as the generalized AUC. The mathematical properties of the measures are analyzed and comparisons between the measures are provided. Two complementary measures are introduced in the thesis, which can be used to support validation conclusions. In the third chapter, a simulation is presented that illustrates a situation where most of the observations in the validation sample carry relatively low losses and a comparison across the measures defined in the second chapter is provided.

CERCS research specialisation: P160 Statistics, operations research, programming, financial and actuarial mathematics.

Keywords. Credit risk, Basel Accords, internal ratings-based models, loss given default models, discriminatory power, rank correlation, generalized AUC.

Table of Contents

Abbreviations

CCF - credit conversion factor

CRR - Capital Requirements Regulation, 575/2013/EU

EAD - exposure at default

EBA - European Banking Authority

ECB - European Central Bank

IRB - internal ratings-based

PD - probability of default

LGD - loss given default

LTV - loan-to-value (collateral coverage)

RWA - risk-weighted assets

Introduction

Financial institutions are exposed to several types of different risks. Credit risk arises from borrowers, who do not meet their contractual obligations. In order to mitigate the risk, financial institutions are required to develop and maintain a framework that quantifies unexpected losses and ensures that they hold sufficient funds to cover for those losses. The foundations of the framework are defined by the Basel Accords, which have been developed by the Basel Committee on Banking Supervision and adopted across the European Union.

For the quantification of credit risk, the Basel Accords introduce a measure called the risk-weighted assets. The volume of risk-weighted assets across portfolios is incorporated into regulatory capital calculations and has to be reported to external regulators on a frequent basis. If a financial institution is compliant with the requirements for the internal ratings-based approach (described in the Basel Accords), they may be permitted to use their own internal risk models to assess the intermediate risk parameters (such as probability of default and loss given default) for the calculation of risk-weighted assets. However, the requirements to obtain the permission are extensive: among other things, the application for internal ratings-based approach assumes that the institution has sufficient internal data for adequate and stable risk classifications, established internal processes to continuously monitor the performance of the models and internal policies that govern the annual validation of the developed models. A general overview of the Basel Accords and the requirements related to regulatory capital calculations are provided in Chapter 1

The objective of the thesis is to examine measures that can be used to assess the discriminatory power of loss given default models. This is a relevant topic related to the validation of internal ratings-based models. Notably, the European Central Bank's instructions for reporting the validation results of internal models include a measure that is referred to as the generalized AUC, which is defined and examined in detail in Chapter 2. In particular, the behaviour of the measure in case of samples with significant concentrations of tied observations is analyzed and suggestions how to support the validation results in such instances are provided. Additionally, alternative measures of discriminatory power are highlighted. In Chapter 3, a simulation is presented that illustrates a situation where most of the observations in the validation sample carry relatively low losses and a comparison across the measures defined in Chapter 2 is provided.

1 Basel Accords and credit risk models

Financial institutions are exposed to several types of different risks, which are mitigated by extensive oversight from internal (risk management departments, internal audit, ...) and external (local government, European Union, ...) regulators. Credit risk arises from borrowers, who do not meet their contractual obligations. Credit risk can be mitigated by accurately quantifying losses that financial institutions observe in case of late or missed payments, taking into account the time-value of money.

This chapter provides a general overview of the Basel Accords, which (among other things) formulate the requirements for the measurement and reporting of various credit risk parameters for financial institutions across the world. As the adoption of requirements set out by the Basel Accords may vary across regions, the thesis focuses on the adoption of Basel accords in the European Union. In addition, the current chapter highlights the general concepts of internal ratings-based credit risk models: their development and internal validation. In particular, the validation of internal ratings-based loss given default models is most relevant for the topic of the thesis.

1.1 General overview of the Basel Accords

The following section is composed based on the descriptions provided by the Bank for International Settlements [1] and the Basel Accords [2] [4] [6], while also considering the regulation (EU) No 575/2013 [22] (referred to as the Capital Requirements Regulation).

The countries in the European Union have agreed on a framework to regulate how much money banks must at least keep in relation to their commitments. The purpose of the adopted framework (i.e. the Basel Accords) is to minimize the risk of problems in the financial system and to dictate how banks should handle risks. The Basel accords were developed and introduced by The Basel Committee on Banking Supervision, which initially comprised the central banks and governors of the G10 countries, but has since expanded and currently consists of 45 members from 28 jurisdictions. The framework regulates different types of risks such as market risk, operational risk and credit risk. The current thesis focuses on credit risk.

1.1.1 The Basel Concordant and Basel I

From the start, one of the main aims of the Basel Committee was to establish and harmonize the supervision of international banking. Their first main publication, the Basel Concordant [11] issued in 1975, set the supervisory principles for banks operating outside their local jurisdiction. The Concordant has been later revised and re-issued (see ”Principles for the supervision of banks’ foreign establishments” [12]). During the 1980s, the capital adequacy of financial institutions became a key discussion point within the Committee, who were concerned of the deterioration of banks' capital ratios during the periods of increased international risks. This set the scene for the development of a unified framework to quantify and assess such risks in relation to financial institutions.

The Basel Accords are made up of three sets of regulations: Basel I, II and III. The first of those, Basel I [2], was published in 1988 and introduced a framework to estimate the minimum ratio of capital in relation to risk-weighted assets (RWA). The classification system proposed by the initial framework mainly addressed credit risk and divided the institution's assets into five risk categories. The institution's assets were to be distributed across those categories based on the types of the commitments. Each category was assigned its own fixed risk weight (0%, 10%, 20%, 50% or 100%, see Annex 2 of [2] for details). For example, the claims on banks incorporated in the OECD were set to be accounted for with a 20% risk-weight whilst the risk-weight for loans fully secured by mortgage on residential property was set at 50%.

The RWA for a particular asset is obtained by multiplying its risk-weight to the exposure value of the financial asset (which in simplified terms represents the amount that the bank would lose in case the debtor is unable to meet its contractual obligations). Exposures are typically split into two components: the on-balance exposure and the off-balance exposure. The on-balance exposure generally represents the sum of the drawn amount (e.g. the loan amount), accrued interest and fees associated with the commitment. The off-balance sheet exposures represent potential but not yet materialized parts of assets (for example: unused part of a credit line). A credit conversion factor (outlined in Annex 3 of [2]) might be applied to off-balance sheet exposures to reduce the RWA.

Institution's capital (which is compared against RWA) was divided into two tiers. Tier 1 capital (referred to as the core capital) consisted of paid-up share capital and disclosed reserves, whilst Tier 2 capital (referred to as the supplementary capital) included undisclosed reserves, general provisions, asset revaluation reserves, subordinated debt and hybrid capital instruments. A detailed overview of capital that was to be included in the capital base is provided in Annex 1 of [2]. The total capital held by the institution (i.e. the sum of Tier 1 and Tier 2 capital) was set to be at least 8% of the summarized RWA across its eligible assets. In addition, the sum of Tier 2 capital was required not to exceed the total Tier 1 capital. In 1996, the Basel Committee released "Amendment to the capital accord to incorporate market risk" [3] which defines a third tier of capital specific to market risk together with the initial set of minimum capital requirements for market risk.

1.1.2 Basel II

After the publication of the first set of Basel regulations, the framework was gradually refined to enhance the credit risk mitigation framework and to address risks other than credit risk. The revised framework, Basel II [4], published in 2004 and revised in 2005, is built upon three pillars:

- Pillar 1: minimum capital requirement,
- Pillar 2: supervisory review,
- Pillar 3: market discipline.

The first pillar defines the minimum capital requirements and quantifies regulatory capital calculations across three major risk dimensions: market risk, operational risk and credit risk. The second pillar can be treated as a rulebook for supervisors to review institutions' compliance to the framework. The third pillar addresses the disclosure requirements for institutions to enhance corporate governance and to unify capital adequacy reporting. The first Basel framework only addressed parts of each of those pillars.

Under the new regulation, the total RWA is calculated by multiplying the capital requirements for operational risk and market risk (outlined in Part 2, Sections V and VI of [4] respectively) by 12.5 and adding the results to the sum of RWA for credit risk. The coefficient 12.5 represents the reciprocal of 8%, i.e. the minimum capital ratio. Additionally, for the measurement of Pillar I credit risk, an expanded concept of RWA was introduced with Basel II that proposes two main estimation approaches for institutions which are examined in more detail below: the standardized approach and the internal ratings-based (IRB) approach.

Before proceeding with details, it must be noted that even though the general framework and a large portion of the requirements published with the second iteration of the Basel Accords are still relevant, those requirements have since been refined and outlined in greater detail across directives and guidelines that have been published by various regulatory institutions (e.g. the European Central Bank and the European Banking Authority in the European Union). Further references and specifications related to the adoption of the Basel Accords in the European Union with regard to credit risk are provided in Section 1.2.

The standardized approach (described in Part 2, Section II of [4]) is a revised version of the RWA calculation according to Basel I, where the differences across various asset classes are outlined in much greater detail and external ratings from recognized sources are incorporated into the risk-assessment process. The regulation also explicitly defines the rules for credit risk mitigation techniques, which are to be considered when applying either of the two approaches (standardized or IRB). The credit risk mitigation techniques allow to reduce the risk-weights for assets that have eligible financial collaterals linked to them. However, the rules for being classified as an eligible financial collateral are quite strict (see Part 2, Section II.D of [4] for further details). Under the standardized approach, exposures are risk-weighted net of specific provisions.

The use of internal ratings-based approach for determining the capital requirements requires supervisory approval and is subject to strict set of minimum requirements that need to be fulfilled. The calculation of capital requirements under IRB approach relies on internal estimates of various risk components such as the probability of default (PD), loss given default (LGD), exposure at default (EAD) and the effective maturity. Those components are used to measure the unexpected losses (which stem from the risk-weight function) and the expected losses. Note that the expected losses defined in the IRB regulation should not be confused with the Expected Credit Losses defined by the IFRS 9 accounting standard and relate to provisioning (discussed in more detail in Section 1.1.3). Institutions applying the IRB approach should risk-weight exposures gross of specific provisions (see paragraph 308 of [4]).

Under the IRB approach, institutions should distribute exposures across five asset classes which provide an initial grouping of underlying risk characteristics: corporate, sovereign, bank, retail and equity. The classification principles are outlined in Part 2, Section III.B of [4]. Generally, exposures to individuals (such as personal term loans, leases, credit cards, residential mortgage loans) are classified as retail exposures. Loans that are extended to small businesses and managed as retail exposures can also be treated as retail exposures, given that total exposure to the borrower group is less than 1 million €.

There are two types of IRB approaches: the foundation IRB and the advanced IRB. Using the foundation approach, the institution is permitted to calculate only their own PD parameter (other parameters are given standardized values). This approach is not available for retail exposures. Under the advanced approach, an institution can assess all of the RW parameters (PD, LGD, EAD, maturity) internally. One of the most important definitions for the estimation of risk components (PD, LGD, EAD) is the definition of default, which is outlined across paragraphs 452-457 of [4] for the IRB approach. In simplified terms, a default event is considered to have occurred with regard to a particular obligor when either of the following conditions are met: the obligor is past due more than 90 days on any material credit obligation; or the institution considers that the obligor is unlikely to pay its credit obligations in full. The indications of unlikeliness to pay are highlighted in paragraph 453 of [4] and include putting the credit obligation to non-accrued status, selling the credit obligation at a material credit-related economic loss, a distressed restructuring of the credit obligation and obligor's bankruptcy.

The expected loss for corporate, sovereign, bank, and retail exposures that are not in default is to be estimated as the product of PD, LGD and EAD. For defaulted exposures, institutions must use their best estimate of expected loss (as outlined in paragraph 471 of [4]) for the measurement. In simplified terms, the expected loss amounts are used to adjust the institution's capital that will be compared to the estimated unexpected loss amount, i.e. the sum of RWA (see paragraph 43 of [4] for further details).

The calculation of unexpected loss is more complex and relies on the risk-weight formula that is specified separately for exposures to corporates, sovereigns and banks (see Figure 1.1 for illustration) and for different types of retail exposures (see Figure 1.2 for illustration). A standalone specification is also made for equity exposures and purchased receivables (see Part 2, Section III.E and III.F of [4] for further details) which are not examined in detail in scope of the current thesis.

272. Throughout this section, PD and LGD are measured as decimals, and EAD is measured as currency (e.g. euros), except where explicitly noted otherwise. For exposures not in default, the formula for calculating risk-weighted assets is:[66, 67]

Correlation (R) = 0.12 × (1 – EXP (-50 × PD)) / (1 – EXP (-50)) +
0.24 × [1 - (1 - EXP(-50 × PD))/(1 - EXP(-50))]

Maturity adjustment (b) = (0.11852 – 0.05478 × ln (PD))^2

Capital requirement[68] (K) = [LGD × N [(1 - R)^-0.5 × G (PD) + (R / (1 - R))^0.5 × G (0.999)] – PD x LGD] x (1 - 1.5 x b)^ -1 × (1 + (M - 2.5) × b)

Risk-weighted assets (RWA) = K x 12.5 x EAD

The capital requirement (K) for a defaulted exposure is equal to the greater of zero and the difference between its LGD (described in paragraph 468) and the bank's best estimate of expected loss (described in paragraph 471). The risk-weighted asset amount for the defaulted exposure is the product of K, 12.5, and the EAD.

[66] Ln denotes the natural logarithm.

[67] N (x) denotes the cumulative distribution function for a standard normal random variable (i.e. the probability that a normal random variable with mean zero and variance of one is less than or equal to x). G (z) denotes the inverse cumulative distribution function for a standard normal random variable (i.e. the value of x such that N(x) = z). The normal cumulative distribution function and the inverse of the normal cumulative distribution function are, for example, available in Excel as the functions NORMSDIST and NORMSINV.

[68] If this calculation results in a negative capital charge for any individual sovereign exposure, banks should apply a zero capital charge for that exposure.

Figure 1.1: The RWA formula for exposures to corporates, sovereigns and banks, an outtake from paragraph 272 of Basel II

Correlation (R) = 0.03 × (1 - EXP(-35 × PD)) / (1 - EXP(-35)) +
0.16 × [1 - (1 - EXP(-35 × PD))/(1 - EXP(-35))]

Capital requirement (K) = LGD × N[(1 - R)^-0.5 × G(PD) + (R / (1 - R))^0.5 × G(0.999)]
– PD x LGD

Risk-weighted assets = K x 12.5 x EAD

The capital requirement (K) for a defaulted exposure is equal to the greater of zero and the difference between its LGD (described in paragraph 468) and the bank's best estimate of expected loss (described in paragraph 471). The risk-weighted asset amount for the defaulted exposure is the product of K, 12.5, and the EAD.

Figure 1.2: The RWA formula for retail exposures, an outtake from paragraph 330 of Basel II

The figures above are meant to serve as an illustration of the RWA formulas presented in the framework and should not be interpreted as definitive rules for the calculation as there are many exceptions and specifications made across the guideline. For instance, the correlation R in the formula for retail exposures that are not in default is set to 0.15 for residential mortgage exposures and to 0.04 for qualifying revolving exposures. The capital requirement K multiplied by 12.5 (which represents the reciprocal of 8%) can be interpreted as the risk-weight. In addition, a scaling factor (currently set at 1.06 for banks operating in the European Union, see Articles

153 and 154 of [22]) might be applied to the risk-weight formula.

Part 2, Section III.H of [4] outlines the minimum requirements for institutions to apply for the IRB approach and covers topics such as the rating system design, risk rating system operations, use of internal ratings, risk quantification, validation of internal estimates, corporate governance and disclosure requirements. In case an institution with IRB permission becomes non-compliant, the supervisors may reconsider its eligibility for the IRB approach.

The collection of methods, processes, controls and IT systems involved in the operation of data collection, risk quantification and the assignment of internal risk estimates is referred to as the rating system. Each rating system should assess the risks considering two distinct dimensions: the risk of borrower default and the risks related to transaction-specific factors. In case of corporate, sovereign and bank exposures, the exposures to the same borrower must all be assigned to the same borrower grade. A borrower grade represents a distinct set of rating criteria, from which the PD estimates are derived. In simplified terms, this means that all the credit contracts assigned to a particular counterparty (who is treated as a corporate, a sovereign or a bank) should obtain an identical PD estimate that reflects the risk grade of the obligor. Each institution is responsible of developing a PD estimation framework that provides a meaningful distribution of exposures and borrowers across grades, without excessive and undue concentrations. As a minimum, each institution is required to set-up seven borrower grades for non-defaulted customers and one for defaulted customers. The transaction-specific factors include the type of the product and the existence of a collateral. Those factors are to be considered when estimating the LGD, which is generally done on facility level (i.e. on contract level).

Retail exposures should be classified into homogeneous pools which provide sufficient differentiation of risk whilst enabling accurate and consistent estimation of loss characteristics at pool level. Institutions should consider both the borrower risk characteristics and the transaction-specific factors when assigning exposures to pools. One example of such a pool could be the distinction of retail mortgage loans. The estimation of PD for retail exposures is thus performed on pool level, rather than on obligor level. However, as stated above, obligor-specific characteristics should still be considered when estimating the default risk for individual facilities. For instance, if an individual already has an active mortgage loan, this information should be considered when issuing a new term-loan on top of the mortgage. Besides PD, institutions should also assess the LGD and EAD for each pool. However, it is not prohibited for multiple pools to share identical PD, LGD and EAD estimates. Retail exposures are generally also assigned risk grades to further differentiate between facilities across- and inside the pools. As with corporate, sovereign and bank exposures, the assignment to risk grades relates to PD estimation.

The model development process can rely on both internal- and external data which is representative of the long run experience and relevant to current and foreseeable conditions. The estimates of PD typically represent the long-run average of one-year default rates. For retail exposures, the estimates of PD can also be derived based on the estimated long-run loss rate and appropriate estimates of LGD (see paragraph 465 of [4] for details). In all cases, the length of the historical observation period for PD model development should be at least five years. If data is available across a longer observation-horizon and deemed relevant, a longer time period should be used for model development. The PD for defaulted exposures should be set to 100%. Details regarding risk quantification under the IRB approach and requirements related to PD, LGD and EAD model development are highlighted in Part 2, Section III.H of [4]. From a practical perspective, the internally developed PD models are usually regression models that are built up using statistically significant obligor- and / or transaction specific variables. One of the main challenges for model developers is the collection and treatment of the input variables, considering that the used data needs to be consistent, relevant to the latest market practices and also available across a sufficient time-horizon.

For institutions applying the advanced IRB approach, the LGD estimates should reflect economic downturn conditions and need to be above the default-weighted average loss rate (given default) which is calculated based on the economic loss of all available defaults in the data source for that type of facility. All the relevant factors should be considered when measuring economic loss, which include material discount effects, material direct- and indirect costs that are associated with collecting on the instrument and the institution's own workout costs. For defaulted exposures, institutions are also required to construct their best estimate of expected loss which is based on the current economic circumstances and facility status. The data used for LGD model development should ideally cover a complete economic cycle, but the observation-horizon must not be shorter than 5 years for retail exposures and 7 years for exposures to corporates, sovereigns and banks. The less data there is, the more conservative an institution must be in its estimations. Additional references to the requirements on LGD estimation for institutions operating in the European Union are provided in Section 1.2.

From a practical perspective, one of the challenges for LGD model developers is the identification of risk drivers, specially for models that are applied on performing (non-defaulted) portfolio. For agreements with pledged collaterals (e.g. mortgage loans), a commonly used tool for distinguishing between high- and low losses is the loan-to-value (LTV), which generally represents the loan amount divided by the collateral's market value. By selling the underlying collateral, the lender can recover a portion of the loan amount that would otherwise be lost. Naturally, the recovered portion varies across different collateral types. The predictive power of the developed model depends strongly on the quality and the treatment of data. As the re-evaluation of collaterals is a time-consuming and rather expensive process, the estimation of market values typically relies on indexation (e.g. on the average market prices). If the actual market prices of collaterals are not accurately captured by indexation or not representative of the portfolio they are applied on, the fitted LGD model will likely contain a strong bias. Even if the market values have been assessed correctly at the time of default, the realization of the collateral can take time and a long-term drop in average market prices after default could mean that the collateral can only be sold for a reduced price. In those instances, it is sometimes reasonable to wait until the market recovers and sell the collateral when the prices are back up. However, this can still result in high loss rates when measured against the default time exposure, as the recovery cash-flows are typically discounted back to the time of default in model development. For agreements with no pledged collaterals, the estimation process can rely on the type of product or the type of customer / agreement. Various clustering methods such as decision trees can be applied to differentiate between the losses across those categorical variables.

EAD parameter represents the gross exposure of the facility upon default of the obligor. For on-balance sheet items, the estimated EAD should not be below the drawn loan amount at the time of reporting. For off-balance sheet items, the approximation of EAD typically relies on the estimation of the credit conversion factor (CCF) and the unused limit amount. In simplified terms, CCF represents the proportion of the unused limit that is estimated to be drawn by the time the facility defaults. Institutions applying the advanced IRB approach must estimate EAD as the long-run default-weighted average across similar facilities and borrowers. Additionally, the volatility of EAD across the economic cycle should be examined and the estimates that are appropriate for an economic downturn must be applied if those are higher than the long-run average EAD. Similarly to LGD, the data used for EAD model development should ideally cover a complete economic cycle, but the observation-horizon must not be shorter than 5 years for retail exposures and 7 years for exposures to corporates, sovereigns and banks. From a practical perspective, the estimates of CCF are typically modelled by observing the behaviour of historically defaulted facilities across the months leading up to the default event. For the grouping of product- or facility groups with similar behaviour, various clustering methods can be used.

Institutions need to outline the rules and processes for grading / pooling of exposures in detail and document their rating systems' design together with operational details. This enables those in charge of rating assignment to carry out the process in a consistent manner and independent third parties (such as internal audit) to clearly understand the rules related to rating assignment. During loan-approval process, each exposure needs to be assigned an appropriate risk ratings and periodic rating reviews must be in place. Institutions must also collect and store the data on borrower- and facility attributes. The stored data should enable to retrospectively allocate exposures to grades and pools. Institutions with the advanced IRB permission should also retain data related to EAD and LGD estimation together with the realized outcomes on LGD and EAD for defaulted facilities. Additionally, IRB models are subject to stress testing, which in simplified terms is a process that focuses on identifying future events that might lead to unfavourable changes on the institution's credit portfolio (such as economic recession) and assesses the ability to withstand changes related to such events with regard to regulatory capital requirements. As an example, an institution could consider a scenario where the GDP decreases for two consecutive quarters and assess the effects on its PD, LGD and EAD levels (for example by anticipating rating migrations and the decrease in average collateral market values).

Each institution must establish independent credit risk control units that are responsible for the design, performance and implementation of the rating system. The credit risk control units should (among other things) test and monitor internal ratings, produce and analyze reports on the rating system, verify that the rating definitions are applied in a consistent manner and document changes to the rating process together with the reasons for the changes. On top of this, a periodic review of the rating system should be performed by the internal validation function. Each institution that employs internal models is to establish well-articulated model validation standards and a robust system to assess the accuracy and consistency of the models. The validation process consists of both quantitative and qualitative aspects, which include the evaluation of the actual performance of the models against the expected performance, comparisons with external data sources, assessment on model documentation and design and a review of the model implementation process. Further details on the model validation function are highlighted in Section 1.2.

1.1.3 Basel III and its relation to IFRS 9 accounting standards

In light of the 2007-2009 financial crisis, it was evident that there was a need to revise and strengthen the capital requirements defined by the first two sets of Basel Accords. The main concerns were related to poor risk management and inadequate liquidity- and credit buffers. The reform package, referred to as Basel III, was published in December 2010 in two parts. The first part, ”Basel III: International framework for liquidity risk measurement, standards and monitoring” [5], was aimed at the management of liquidity risk and was later (in 2013 and 2014) replaced by ”Basel III: The Liquidity Coverage Ratio and liquidity risk monitoring tools” [7] and ”Basel III: the net stable funding ratio” [8]. The second part, ”Basel III: A global regulatory framework for more resilient banks and banking systems" [6], focused on enhancing the capital requirements set out by the Basel I and Basel II frameworks. Some notable changes proposed by [6] on the credit risk framework are discussed below. Additionally a reference to the revision of the Basel III framework [9] (published in 2017 and unofficially referred to as Basel IV) is provided.

The revised minimum capital requirements are outlined in Part 1 of [6]. Paragraphs 49-96 highlight the updated definitions and requirements on internal capital, including the split of Tier 1 capital into Common Equity Tier 1 (CET1) capital and Additional Tier 1 capital. Notably, the total Tier 1 capital and the total CET1 capital are required to amount to at least 6% and 4.5% of the institution's RWA respectively. The 8% requirement on the sum of Tier 1 and Tier 2 capital (proposed by [2] and adopted by [4]) was not changed. Paragraphs 122-135 set out the requirements for institutions to hold an additional capital buffer on top of the minimum capital requirement. In simplified terms, the aim of the capital conservation buffer is to progressively limit the distribution of earnings (such as dividend payments and bonus payments to staff) in case the capital ratio approaches the regulatory minimum level. On top of this, local regulators may impose a national countercyclical buffer during periods of excess credit growth to counter system-wide risks related to potential future losses (see paragraphs 136-150 of [6] for details). Paragraphs 97-121 highlight various additional amendments to the Basel II framework. For instance, according to paragraph 102, a multiplier of 1.25 should be applied on the correlation parameter in the risk-weight formula (see Figure 1.1) for regulated financial institutions whose total assets amount to 100 billion $ or more.

Additionally, paragraphs 23-25 of [6] discuss the revision of provisioning practices by the International Accounting Standards Board. In simplified terms, provisions represent liabilities in institution's balance sheet that are set aside to anticipate future losses. The Basel Committee on Banking Supervision expressed its support for moving from the provisioning approach that is based on incurred losses (i.e. is backwards-looking) to an approach that is based on estimating expected credit losses (i.e. is forward-looking). In 2018, the accounting standard "IAS39 Financial Instruments: Recognition and Measurement" [26] was in most parts (e.g. with an exception for hedge accounting) replaced by the standard "IFRS 9 Financial Instruments" [27], which was first published in 2014 and describes the revised provisioning principles. To an extent, the estimation of expected credit losses (ECL) defined by the IFRS 9 standard bears a resemblance to the estimation of expected losses according to the Basel Accords. Typically (although not explicitly required by the standard), the estimation of ECL is also based on the calculation of various intermediate risk components such as PD, LGD and EAD, which are subject to periodic review. Thus, the methods outlined in Chapter 2 for the measurement of discriminatory power can also be applied on IFRS 9 LGD (and ECL) models. However, as opposed to the measurement of expected losses under IRB approach, ECL is aimed to be precise (not stable / conservative) and should follow the changes in credit cycle, considering different economic scenarios and forecasts. In addition, ECL has to be estimated both for the following 1-year period as well as for the expected lifetime of the financial instrument. The estimation-horizons for a particular facility is determined by comparing the estimated risk level at reporting date to the risk assessment made at the initial recognition of the commitment. If the risk levels have significantly increased since initial recognition or if the facility is credit-impaired (e.g. defaulted), then lifetime ECL has to be estimated.

In 2017, The Basel Committee on Banking Supervision published the document "Basel III: Finalising post-crisis reforms" [9], which is unofficially referred to as Basel IV. The main aim of the revision is to reduce the variability in RWA calculations across financial institutions. Among other things, the framework should increase the risk sensitivity of standardized risk estimation approaches while constraining approaches that are based on internal models. The general principles of IRB framework that were described in Section 1.1.2 are still relevant in the revised framework (e.g. the RWA formulas provided in Figures 1.1 and 1.2), but additional requirements and limitations are proposed. For example, constraints are set on the use of internal models by introducing various output floors to estimated risk metrics. For institutions with IRB permission, the final RWA amount that will be measured against internal capital should not be lower than 72.5% of the RWA that is calculated using the standardized approach (even if internal models suggest lower RWA). This implies that institutions with IRB permission also need to estimate the risk using the standardized approach in parallel. Additionally, minimum levels for internally estimated LGD parameters were set (which depend on the type of customer / collateral) and minimum PD levels were revised (in generalized terms, the PD floor was increased from 0.03% to 0.05%). The framework was set to be implemented by the start of 2022, but the adoption date has since been postponed in light of the global pandemic.

1.2 Adoption of the Basel Accords in the European Union

This section provides a brief overview of the laws and regulatory guidelines that have been adopted in the European Union with respect to the Basel accords and IRB models.

The Capital Requirements Regulation (CRR) [22] that was published in 2013 and has been adopted since 2014 is a legislative act that reflects the capital requirements set out by the Basel Accords (Basel I, II and III) for financial institutions operating in the European Union. The regulation specifies the aspects of credit risk assessments and capital requirements outlined in Section 1.1 that had been formulated by the Basel Committee at the time of the publication. Namely, Part II of [22] provides the definitions of own funds including the definitions of Tier 1 and Tier 2 capital, Title I of Part III sets the own funds requirements for financial institutions and includes the 4.5%, 6% and 8% requirements on Tier 1 and Tier 2 capital that were highlighted in Section 1.1.3, Title II of Part III covers the capital requirements related to credit risk and includes detailed descriptions of the standardized approach and the IRB approach with all the related definitions such as the definition of default, exposure, risk components and the risk-weights. In 2019, a reform package to CRR was published in the form of Regulation (EU) No 2019/876 [23] which is referred to as CRR2. In terms of credit risk, one of the most notable changes proposed by CRR2 is the reduction of RWA for small- and medium sized enterprises with exposure above 1.5 million €. Additionally, the revision addresses topics such as the leverage ratio, the net stable funding ratio, market risk and disclosure requirements. For the most part, the requirements set out in CRR2 were to be implemented by June 2021. In October 2021, the European Commission released its proposal for amending Regulation (EU) No 575/2013 with regard to requirements for credit risk, credit valuation adjustment risk, operational risk, market risk and the output floor [24], which is referred to as CRR3 and marks the adoption of ”Basel III: Finalising post-crisis reforms” (i.e. the requirements of Basel IV). In most parts, the proposed regulation is set to be applied starting from 2025.

Besides the regulatory capital requirements, the European Banking Authority (EBA) and the European Central Bank (ECB) have published several guidelines and standards related to internal models. In July 2016, EBA published the ”Final Draft Regulatory Technical Standards on the specification of the assessment methodology for competent authorities regarding compliance of an institution with the requirements to use IRB Approach" [16]. Although the specifications provided in the standard are quite general, they can still be used as a blueprint for institutions to comply with IRB requirements. Notably, Chapter 3 of [16] discusses the revision of the internal validation function. The standard requests to verify the independence of the validation function, the frequency and completeness of the validation process, the adequacy of methods and procedures that are applied in the validation process and the soundness of the validation reporting process. In particular, it is highlighted that the validation function should be independent from the functions responsible for the model design and development, and the validation function should report directly to senior management. It is also expected that the validation function reviews the procedures related to data collection, the choices of model development methodology, the model parameters implemented in IT systems and the performance of the models with regard to risk quantification, risk differentiation and the stability of internal ratings. Generally, the validation process is expected to be performed at least once a year (i.e. on an annual basis).

The "Draft Regulatory Technical Standards on the materiality threshold for credit obligations past due" [14] and the "Guidelines on the application of the definition of default" [15], published by EBA during 2016 and applied from 2021, were issued to harmonize the default definition across institutions. The guidelines provide clarifications on topics such as the days past due criterion, indications of unlikeliness to pay, return to non-defaulted status and are generally applicable for both the standardized- and the IRB approach. According to [14], the absolute threshold for recognizing days past due must not be higher than 100 € for retail exposures or 500 € for other exposures. Article 59 of [15] outlines additional indications of unlikeliness to pay that can be considered by institutions, which include situations where a borrower's sources of income are no longer available to meet the payments or there are justified concerns about the ability to generate stable and sufficient cash flows, the borrower has breached the agreed terms of a credit contract or the institution has requested additional collateral. Articles 71-74 of [15] highlight the minimum conditions for a default to be reclassified to non-defaulted status and, among other things, state that at least 3 months must pass since the moment that default indicators cease to be met for the reclassification to occur.

The "Guidelines on PD estimation, LGD estimation and the treatment of defaulted exposures" [17] was published by EBA in 2017 and focuses on the modelling approaches for the development of IRB PD and LGD parameters. The guidelines cover a vast variety of topics such as key definitions, data requirements, selection of risk drivers, estimation methodologies, model calibration and the review of developed estimates. Chapters 6 and 7 of [17] discuss the estimation of LGD for non-defaulted and defaulted facilities respectively. Generally, the initial calibration of LGD models revolves around the calculation of long-run average LGD, which represents the average loss of all defaults for a particular grade or pool in the model development sample. For facilities that are in default, two different models need to be developed: a model to estimate the LGD in-default and a model that estimates the expected loss best estimate (as mentioned in Section 1.1.2). Compared to the LGD models that are applied on non-defaulted portfolio, both the LGD in-default model and the expected loss best estimate model should take into consideration the relevant post-default information, for example the time spent in-default. The best estimates of expected loss should be estimated based on the long-run average LGD, but also account for the economic conditions at the time of estimation. However, as the LGD that is used for regulatory capital calculations needs to be appropriate for an economic downturn (if those estimates are more conservative than the long-run average, see Article 181 of [22]), then both the LGD model for non-defaulted facilities and the LGD in-default model need to be typically further calibrated to be used in the RWA calculation. For this purpose EBA published two additional guidelines: the ”Final Draft Regulatory Technical Standards on the specification of the nature, severity and duration of an economic downturn” [18] and the ”Final report on guidelines for the estimation of LGD appropriate for an economic downturn” [19]. Notably, three different downturn LGD estimation approaches are outlined in [19]: an approach based on observed impact that depends on historic loss data, an approach based on estimated impact that incorporates extrapolation of historic loss data and is used in situations where sufficient data on downturn conditions is not available, and a third approach that is used in situations where neither of the first two approaches are applicable and incorporates a 15% add-on to the long-run average LGD estimates. All of the three guidelines mentioned in this paragraph ([17], [18] and [19]) are applied since the start of 2021.

In 2019, ECB published the consolidated version of their guide to internal models [20]. The document reflects ECB's views and interpretation of the CRR, provides a summarized overview across the key articles of CRR and (among other things) includes relevant references to the guidelines / standards mentioned in this section. The guideline is split into 4 chapters: general topics, credit risk, market risk and counterparty credit risk. Notably, Section 4 of the general topics chapter discusses internal validation and outlines the key areas of independent validation routine: back-testing, measurement of discriminatory power (for PD, LGD and CCF models), analyses of representativeness, analyses of overrides, stability analyses of internal ratings and risk parameters, analyses of model design, evaluation of input data, benchmarking analyses, data cleansing analyses, review of model specification and the quality assurance of the used computer codes. Additionally, it is highlighted that quantitative thresholds should be set up for the assessment of the back-test results, measurement of discriminatory power, analyses of overrides and the stability of the internal ratings / risk parameters. Areas where no quantitative thresholds are applied should adhere to consistent qualitative assessments. During the same year, ECB also published instructions for reporting the validation results of IRB credit risk models [21], which are examined in more detail in Section 1.2.1 below.

1.2.1 ECB instructions for reporting the validation results of internal models

After an institution has obtained the permission to use IRB credit risk models for regulatory capital calculations, their compliance with the regulation is monitored on a regular basis. An important aspect of the supervision is the review of internal validation results, which should be shared with the ECB within one month after the validation results have been finalized and internally approved [21]. As the validation routines may vary across institutions (e.g. the statistical test that are performed), supplementary validation reporting is requested by the ECB which enables to compare the models used by institutions on a common basis.

Majority of the validation techniques described in [21] relate to quantitative aspects of internal models. The document highlights distinct sets of requirements for reporting the results across IRB risk components, including PD, LGD, expected loss best estimate, LGD in-default and CCF. Section 2.4 of [21] details the general information that should be provided on the models, including an overview of the application portfolio (e.g. the total portfolio exposure, number of facilities, number of customers) and the internal validation assessment on the model. Section 2.5 describes the reporting requirements on PD models, which include analyses on the predictive ability, discriminatory power and stability of the model. Notably, the predictive ability of PD models is assessed using a Jeffrey's test (see Section 2.5.3 of [21] for details) and the discriminatory power of PD models is measured by calculating the area under the ROC curve (see

Section 2.5.4 and Annex 3.1 of [21] for details).

Sections 2.6-2.9 of [21] highlight the reporting requirements for LGD, expected loss best estimate, LGD in-default and CCF models. For the assessment of predictive ability, a one sample t-test for paired observations is used for all four types of models. The t-tests that should be performed differ by the parameters the test is applied on, but also by the observation horizon and population selection criteria. For further details, see Sections 2.6.2, 2.7.2, 2.8.2 and 2.9.3 of [21]. Additionally, discriminatory power should be measured for both the LGD models that are applied on the non-defaulted portfolio and the CCF models. The assessment of discriminatory power is based on the calculation of measure referred to as the generalized AUC, which is defined in Sections 2.6.3, 2.9.4 and Annex 3.2 of [21]. The measure is examined in detail in Chapter 2 of the current thesis.

In order to achieve consistency in the supplementary validation reporting, templates that highlight the required information for each model are submitted by the ECB to institutions with IRB permission. Besides general information, the templates should be filled with statistics on the samples constructed for quantitative testing purposes (typically in the form of a contingency table). The statistical tests required by [21] are integrated into the templates and if filled correctly, the results of the tests are provided automatically (without the need to input the formulas manually).

2 Measures to assess the discriminatory power of loss given default models

Paragraph 2.6 of [21] outlines the requirements for reporting the validation results of Internal ratings-based loss given default models. Among other things, the institutions with IRB permission are required to perform the Generalized AUC (gAUC) test on model validation data and present the results according to the requirements of Section 2.6.3 and Annex 3.2 of [21]. In the following, we are going to define the measure and explore its mathematical properties with regard to LGD models and the regulatory reporting template.

2.1 Kendall's rank correlation coefficient

Kendall's rank correlation coefficient is a measure of ordinal association between two quantities. It was proposed by Maurice George Kendall in his 1938 paper "A new measure of rank correlation" [29]. Definitions 2.2 and 2.4 presented in Section 2.1.1, treated as prerequisites for the measurement of gAUC, are derived based on that publication.

2.1.1 Kendall's rank correlation coefficient for given samples

2.1.1.1 Kendall's rank correlation coefficient for one-dimensional vectors

Consider a population of n entities and let $\mathbf{x} = (x_1, ..., x_n)$ be a vector of observations, for example the predictions of a model sorted by the observed realizations in case of a particular portfolio on a bank's balance sheet ($n \in \mathbb{N}, n \geq 2, x_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$). Let x_i and x_j be two random observations from the vector, $j > i$. The observations x_i and x_j are called *concordant* if $x_j > x_i$. If $x_i < x_j$, then the observations x_i and x_j are called *disconcordant*. If $x_i = x_j$, then the pair is neither concordant nor disconcordant.

Define the *score for vector* $\mathbf{x}$ ($S_{\mathbf{x}}$) as the difference between the number of concordant pairs in the vector and the number of disconcordant pairs in the vector. The calculation of score is performed by comparing the ordering of each pair of observations from the vector. Given two random observations x_i and x_j ($i < j$), if the pair is concordant (i.e. $x_i < x_j$), then the current score (which initially starts from 0) is increased by 1. In case of tied observations ($x_i = x_j$), the order cannot be uniquely determined and the score for the vector is neither increased or decreased. Otherwise, if $x_i > x_j$, then the current score is decreased by 1. Kendall's rank correlation coefficient measures the ratio of the score of the vector and the *maximum possible score for vector x* (denoted by $\overline{S}_{\mathbf{x}}$). The maximum possible score represents the case where all of the pairs in the vector are concordant and is equal to the total number of pairs in the vector.

One way to calculate the score for a vector is to start by comparing the first element x_1 to all succeeding elements and then continue the same process with the second element, until all the pairs have been compared in terms of rank-ordering once. This way, the score for the vector can be calculated by performing $n-1$ comparisons during the first step, $n-2$ comparisons during the second step, up until the single comparison of the last two elements. This amounts to

$$\sum_{i=1}^{n-1} i = \frac{(n-1)\cdot n}{2} = \binom{n}{2},$$

steps, which means that

$$\overline{S}_{\mathbf{x}} = \binom{n}{2}. \tag{2.1}$$

The score for vector $\mathbf{x}$ can be expressed as:

$$S_{\mathbf{x}} = \sum_{i=1}^{n-1}\sum_{j>i} \operatorname{sgn}(x_j - x_i).$$

Corollary 2.1. *Let* $\mathbf{x} = (x_1, ..., x_n)$ *be a vector of observations,* $n \in \mathbb{N}$*,* $n \geq 2$*,* $x_i \in \mathbb{R}$ $\forall i \in \{1, ..., n\}$*. Then*

$$S_{\mathbf{x}} = \overline{S}_{\mathbf{x}} \iff \forall i, j \in 1, ..., n : i < j \iff x_i < x_j.$$

Now it is possible to formulate the definition of Kendall's rank correlation coefficient for the one-dimensional case.

Definition 2.2. *Let* $\mathbf{x} = (x_1, ..., x_n)$ *be a vector of observations,* $n \in \mathbb{N}$*,* $n \geq 2$*,* $x_i \in \mathbb{R}$ $\forall i \in \{1, ..., n\}$*. The Kendall's rank correlation coefficient for vector* $\mathbf{x}$ *is defined as*

$$\tau_{\mathbf{x}} = \frac{S_{\mathbf{x}}}{\overline{S}_{\mathbf{x}}} = \frac{2}{(n-1)\cdot n} \cdot \sum_{i=1}^{n-1}\sum_{j>i} \operatorname{sgn}(x_j - x_i).$$

Corollary 2.3. *Let* $\mathbf{x} = (x_1, ..., x_n)$ *be a vector of observations,* $n \in \mathbb{N}$*,* $n \geq 2$*,* $x_i \in \mathbb{R}$ $\forall i \in \{1, ..., n\}$*. Then*

(a) $1 \geq \tau_{\mathbf{x}} \geq -1$*,*

(b) $x_1 = x_2 = ... = x_n \implies \tau_{\mathbf{x}} = 0.$

2.1.1.2 Kendall's rank correlation coefficient for two-dimensional samples

Although the calculations performed according to Definition 2.2 technically only require a one-dimensional set of observations as an input, it is important to understand that the uniqueness of $\tau_{\mathbf{x}}$ is determined by the ordering of those outcomes, i.e. it is dependent of the index i. From a practical perspective, the sample might not always be ordered and it would be easier to feed the two sets of quantities (the realization and the variable that determines the ordering) into the calculations directly. Such a method can be mathematically formulated by observing that

$$\sum_{i=1}^{n-1}\sum_{j>i} \operatorname{sgn}(x_j - x_i) = \sum_{i=1}^{n-1}\sum_{j>i} \operatorname{sgn}(x_j - x_i) \cdot \operatorname{sgn}(j - i).$$

The definition of Kendall's rank correlation coefficient can thus be expanded to cover randomly ordered two-dimensional samples $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ in the following way:

Definition 2.4. *Let* $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ *be a sample of observations,* $n \in \mathbb{N}$*,* $n \geq 2$*,* $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$*. The Kendall's rank correlation coefficient for sample* $(\mathbf{x}, \mathbf{y})$ *is defined as*

$$\tau_{\mathbf{x},\mathbf{y}} = \frac{S_{\mathbf{x},\mathbf{y}}}{\bar{S}_{\mathbf{x},\mathbf{y}}} = \frac{2}{(n-1)\cdot n} \cdot \sum_{i=1}^{n-1}\sum_{j>i} \operatorname{sgn}(x_j - x_i) \cdot \operatorname{sgn}(y_j - y_i).$$

The calculation (comparison) principles for two-dimensional (not ordered) samples are similar to the one-dimensional (ordered) case, as reflected by the expressions for the corresponding correlation coefficients. However, there are two main differences. Firstly, in case of two-dimensional samples, both components of each observation are compared and thus, the number of comparisons performed in practice is two times higher than in case of one-dimensional vectors. The second difference arises from situations where duplicates occur on the variable that the observations are ordered upon. If two compared realizations are equal (e.g. zero loss outcome), the score is neither increased or decreased in case of two-dimensional samples. For one-dimensional vectors, the score would still be calculated if $y_i = y_j$, but $x_i \neq x_j$. If the concentration of tied observations in the sample across either variable is insignificant, both approaches could be applicable.

Corollary 2.5. *Let* $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ *be a sample of observations,* $n \in \mathbb{N}$*,* $n \geq 2$*,* $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$*. Then* $\tau_{\mathbf{x},\mathbf{y}} = \tau_{\mathbf{y},\mathbf{x}}$*.*

Observe that

$$\tau_{\mathbf{x},\mathbf{x}} = \frac{2}{(n-1)\cdot n} \cdot \sum_{i=1}^{n-1}\sum_{j>i} [\operatorname{sgn}(x_j - x_i)]^2,$$

represents the proportion of pairwise comparable observations according to $\mathbf{x}$, and similarly

$$\tau_{\mathbf{y},\mathbf{y}} = \frac{2}{(n-1)\cdot n} \cdot \sum_{i=1}^{n-1}\sum_{j>i} [\operatorname{sgn}(y_j - y_i)]^2,$$

represents the proportion of pairwise comparable observations according to $\mathbf{y}$. Note that

$$\begin{cases} \tau_{\mathbf{x},\mathbf{x}} = 0 \iff x_i = x_j\ \forall i, j \in \{1, ..., n\}; \\ \tau_{\mathbf{x},\mathbf{x}} \in (0, 1] \iff \exists\, i \neq j : x_i \neq x_j. \end{cases}$$

Remark. Note that $[\operatorname{sgn}(x_j - x_i)]^2 = \mathrm{I}(x_j \neq x_i)$, where $\mathrm{I}(\cdot)$ represents the indicator function.

Example 2.6. *Let* $(\mathbf{x}, \mathbf{y}) = ((2, 1), (2, 2), (3, 1), (3, 3))$*. Even though all the observations in the sample are distinct, duplicates still occur on component level (for example,* $x_1 = x_2 = 2$*). The total number of comparisons across the sample is six. Firstly, starting from observation* $(2, 1)$ *on the left and comparing it to all the following observations, the score for the sample remains zero after first two comparisons, but is increased by one during the last step, as* $(2, 1)$ *and* $(3, 3)$ *are concordant. Having compared the first observation with all succeeding observations, it can be discarded and the process restarts with the second observation in line. The two comparisons with* $(2, 2)$ *do not have an impact on the final score, as* $(2, 2)$ *and* $(3, 1)$ *are disconcordant, but* $(2, 2)$ *and* $(3, 3)$ *are concordant. This leaves one last comparison to cover the full sample: as* $x_3 = x_4 = 3$*, the final score for the sample remains* 1 *and* $\tau_{\mathbf{x},\mathbf{y}} = \frac{1}{6}$*.*

In terms of credit risk, considering both the regulatory reporting templates [21] as well as requirements on IRB credit risk models in general (such as approximating probability of default across risk grades), it is often required to examine and present the estimates and realizations across predefined pools or segments, rather than on a continuous scale. Thus, contingency tables are commonly used for quantifying and analyzing the proportion of tied observations within a population and giving assessments on model's discriminatory power, stability (e.g. migration matrices), predictive power and conservatism. Furthermore, taking into account that the number of comparisons performed for the calculation of Kendall's τ is equal to $\binom{n}{2}$ if all the observations in the sample as treated as unique (see Equality 2.1), grouping of observations could significantly improve the calculation speed, which could become problematic in case of large sample sizes. Thus, we also introduce the calculation principles for Kendall's τ using contingency tables [10].

In the following, it is assumed that the underlying data has been aggregated and presented in the form of a $M_1 \times M_2$ contingency table, where M_1 represents the number of rows in the contingency table and M_2 the number of columns in the contingency table. A similar approach (using contingency tables) has been detailed in Annex 2 of [21] for the calculation of Somers' D. Let $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ be a sample of observations ($n \in \mathbb{N}$, $n \geq 2$, $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$). Denote the number of observations in cell (i, j) of the contingency table as N_{ij}. For each unique observation situated in cell (i, j) of the contingency table, denote the number concordant observations in the sample by N_{ij}^+ and the number of disconcordant observations in the sample by N_{ij}^-. Observe that

$$N_{ij}^+ = \sum_{k<i} \sum_{l<j} N_{kl} + \sum_{k>i} \sum_{l>j} N_{kl},$$

and

$$N_{ij}^- = \sum_{k>i} \sum_{l<j} N_{kl} + \sum_{k<i} \sum_{l>j} N_{kl}.$$

For each observation in the contingency table, the number of concordant (disconcordant) observations can be thought of as the sum of all observations that are situated to the upper-left (lower-left) or lower-right (upper-right) side of the contingency table. For fixed (i, j), the total score for comparisons situated in that cell can be expressed as

$$N_{ij} \cdot (N_{ij}^+ - N_{ij}^-).$$

If the difference between the number of concordant and disconcordant observations is summarized for each cell in the contingency table, then all the elements are pairwise compared to each other twice, yielding

$$2 \cdot S_{\mathbf{x},\mathbf{y}} = \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij} \cdot (N_{ij}^+ - N_{ij}^-).$$

This allows us to formulate the Kendall's rank correlation coefficient using contingency tables as follows:

$$\tau_{\mathbf{x},\mathbf{y}} = \frac{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij} \cdot (N_{ij}^+ - N_{ij}^-)}{(n-1) \cdot n}. \tag{2.2}$$

Example 2.7. *Let* $(\mathbf{x}, \mathbf{y}) = ((2, 1), (2, 2), (3, 1), (3, 3))$ *(as in Example 2.6). The contingency table for the sample can be presented as follows:*

X \ Y	1	2	3
1	0	0	0
2	1	1	0
3	1	0	1

Table 2.1: Contingency table in case of $(\mathbf{x}, \mathbf{y}) = ((2,1),(2,2),(3,1),(3,3))$

For each observation in cell (i, j), the elements in row i and the elements in column j cannot impact the final score as the ordering cannot be distinctly measured. The sum of all elements in the sample is equal to the sum of all elements in the contingency table. Starting from the first non-zero cell, observe that $N_{21}^{+} = N_{32} + N_{33} = 1$ whilst $N_{21}^{-} = N_{12} + N_{13} = 0$. Thus, when accounting for all the elements in cell $(1, 2)$ (which in this case is just one), the score is increased by $1 \cdot (1 - 0) = 1$. Note that as the cell is situated on one of the borders of the contingency table, e.g. on the left hand side of the table, only elements that are situated diagonally to the upper-right side and lower-right side are considered. If the same principle of adding up concordant and disconcordant observations is carried out throughout the contingency table, each comparison is accounted for twice, yielding $2 \cdot S_{\mathbf{x},\mathbf{y}} = 2$, whilst $\tau_{\mathbf{x},\mathbf{y}} = \frac{2 \cdot S_{\mathbf{x},\mathbf{y}}}{n \cdot (n-1)} = \frac{2}{12} = \frac{1}{6}$, as shown before.

2.1.2 Kendall's rank correlation coefficient for random samples

It is evident from Section 2.1.1 that τ measures the difference of the proportions of concordant and disconcordant observations in the sample, which is also mentioned by Robert H. Somers in his article "A New Asymmetric Measure of Association for Ordinal Variables" [35]. Namely, if N_p^+ denotes the number of concordant pairs in the sample, N_p^- the number of disconcordant pairs in the sample and N_p the total number of pairs in the sample, then

$$\tau_{\mathbf{x},\mathbf{y}} = \frac{N_p^+ - N_p^-}{N_p} = \frac{N_p^+}{N_p} - \frac{N_p^-}{N_p}. \tag{2.3}$$

Assume that the sample is defined through two dependent random variables X and Y. Let (X_1, Y_1) and (X_2, Y_2) be two independent copies of (X, Y). Define event A as "(X_1, Y_1) and (X_2, Y_2) are concordant" and event B as "(X_1, Y_1) and (X_2, Y_2) are disconcordant". Denote

$$p_1 := \mathrm{P}[(X_1, Y_1) \text{ and } (X_2, Y_2) \text{ are concordant}] = \mathrm{P}(A), \tag{2.4}$$

and

$$p_2 := \mathrm{P}[(X_1, Y_1) \text{ and } (X_2, Y_2) \text{ are disconcordant}] = \mathrm{P}(B). \tag{2.5}$$

Based on the interpretation of τ as the difference of two proportions (highlighted by Equality 2.3), define the *Kendall's rank correlation coefficient for random variables X and Y* as

$$\tau_{X,Y} = p_1 - p_2.$$

Observe that

$$p_1 = \mathrm{P}[(X_1 > X_2 \wedge Y_1 > Y_2) \vee (X_1 < X_2 \wedge Y_1 < Y_2)] = \mathrm{P}[(X_1 - X_2) \cdot (Y_1 - Y_2) > 0]$$
$$= \mathrm{P}[\mathrm{sgn}(X_1 - X_2) \cdot \mathrm{sgn}(Y_1 - Y_2) = 1],$$

and similarly

$$p_2 = \mathrm{P}[(X_1 < X_2 \wedge Y_1 > Y_2) \vee (X_1 > X_2 \wedge Y_1 < Y_2)] = \mathrm{P}[(X_1 - X_2) \cdot (Y_1 - Y_2) < 0]$$
$$= \mathrm{P}[\mathrm{sgn}(X_1 - X_2) \cdot \mathrm{sgn}(Y_1 - Y_2) = -1].$$

As the random variable $\mathrm{sgn}(X_1 - X_2) \cdot \mathrm{sgn}(Y_1 - Y_2)$ can only obtain three distinct values:

$$\mathrm{sgn}(X_1 - X_2) \cdot \mathrm{sgn}(Y_1 - Y_2) = \begin{cases} 1, \text{ with probability } p_1; \\ -1, \text{ with probability } p_2; \\ 0, \text{ with probability } 1 - p_1 - p_2; \end{cases}$$

then

$$\mathbb{E}[\mathrm{sgn}(X_1 - X_2) \cdot \mathrm{sgn}(Y_1 - Y_2)] = 1 \cdot p_1 + (-1) \cdot p_2 + 0 \cdot (1 - p_1 - p_2) = p_1 - p_2.$$

Corollary 2.8. *Let (X_1, Y_1) and (X_2, Y_2) be two independent copies of (X, Y), where X and Y are dependent random variables. Then*

$$\tau_{X,Y} = \mathbb{E}[\mathrm{sgn}(X_1 - X_2) \cdot \mathrm{sgn}(Y_1 - Y_2)].$$

2.2 Goodman and Kruskal's gamma

An alternative way to measure ordinal association between two quantities was proposed by Leo Goodman and William Kruskal in their paper "Measures of Association for Cross Classifications" [25], which was published in 1954. The proposed method involves a modification to the calculation of the maximum possible score for the sample, where only non-tied pairs are considered for the calculation of the correlation coefficient, denoted by $\gamma_{\mathbf{x},\mathbf{y}}$.

2.2.1 Gamma for given samples

As only non-tied pairs are considered for the calculation of the correlation coefficient, a theoretical bound of requiring a pair of comparable elements in the sample is set. Thus, in the following, it is assumed that examined samples are *non-trivial*, i.e. $\exists k, l,\ 1 \leq k < l \leq n :$
$x_k \neq x_l \wedge y_k \neq y_l$.

Definition 2.9. *Let $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ be a sample of observations, $n \in \mathbb{N}$, $n \geq 2$, $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$. The Goodman and Kruskal's gamma for sample $(\mathbf{x}, \mathbf{y})$ is defined as*

$$\gamma_{\mathbf{x},\mathbf{y}} = \frac{\sum\limits_{i=1}^{n-1} \sum\limits_{j>i} \mathrm{sgn}(x_j - x_i) \cdot \mathrm{sgn}(y_j - y_i)}{\sum\limits_{i=1}^{n-1} \sum\limits_{j>i} [\mathrm{sgn}(x_j - x_i)]^2 \cdot [\mathrm{sgn}(y_j - y_i)]^2}$$

Corollary 2.10. *Let $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ be a sample of observations, $n \in \mathbb{N}$, $n \geq 2$, $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$. Then $\gamma_{\mathbf{x},\mathbf{y}} = \gamma_{\mathbf{y},\mathbf{x}}$.*

Corollary 2.11. *Let* $\mathbf{x} = (x_1, ..., x_n)$ *be a sample of observations,* $n \in \mathbb{N}$*,* $n \geq 2$*,* $x_i \in \mathbb{R}$ $\forall i \in \{1, ..., n\}$*. Then* $\gamma_{\mathbf{x},\mathbf{x}} = 1$*.*

To formulate the definition of $\gamma_{\mathbf{x},\mathbf{y}}$ for contingency tables [10] (similarly to what was shown in Section 2.1.1), observe that for a fixed cell (i, j) in the contingency table, the total number of comparisons with non-zero outcome across the sample can be expressed as

$$N_{ij} \cdot (N_{ij}^{+} + N_{ij}^{-})$$

and thus

$$\gamma_{\mathbf{x},\mathbf{y}} = \frac{\sum\limits_{i=1}^{M_1} \sum\limits_{j=1}^{M_2} N_{ij} \cdot (N_{ij}^{+} - N_{ij}^{-})}{\sum\limits_{i=1}^{M_1} \sum\limits_{j=1}^{M_2} N_{ij} \cdot (N_{ij}^{+} + N_{ij}^{-})}.$$

When comparing γ to τ (see Equality 2.2), the score calculation principles (i.e. the enumerator in the formula) are identical but the denominators differ. In particular,

$$\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij} \cdot (N_{ij}^{+} + N_{ij}^{-}) \leq (n-1) \cdot n,$$

which allows us to formulate the following Corollary (note that this is not a new discovery by the author, but a known fact derived directly from the definitions).

Corollary 2.12. *Let* $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ *be a sample of observations,* $n \in \mathbb{N}$*,* $n \geq 2$*,* $x_i, y_i \in \mathbb{R}$ $\forall i \in \{1, ..., n\}$*. Then* $1 \geq |\gamma_{\mathbf{x},\mathbf{y}}| \geq |\tau_{\mathbf{x},\mathbf{y}}| \geq 0$*.*

Example 2.13. *Let* $(\mathbf{x}, \mathbf{y}) = ((2, 1), (2, 2), (3, 1), (3, 3))$*. In order to estimate* $\gamma_{\mathbf{x},\mathbf{y}}$*, the sum of both concordant and disconcordant observations has to be calculated, which amounts to (considering double-counting)* $1 + 2 + 1 + 2 = 6$*. In Example 2.6 it was shown that* $2 \cdot S_{\mathbf{x},\mathbf{y}} = 2$*. Thus,* $\gamma_{\mathbf{x},\mathbf{y}} = \frac{2}{6} = \frac{1}{3}$*.*

In practice, the difference of having a correlation coefficient of over 0.3 instead of approximately 0.167 on the same sample can make a significant difference. Thus, it is important to note that the extent of how much the two quantities (γ and τ) differ is determined by the proportion of pairwise-tied observations across the sample. Moreover, the decision in terms of which correlation coefficient is more suitable for the problem at hand is not limited to those borderline cases (counting all possible comparisons vs. comparing only those that can make a difference). This is further elaborated upon in Section 2.3.

2.2.2 Gamma for random samples

Let N_p^{+} denote the number of concordant pairs in the sample, N_p^{-} the number of disconcordant pairs in the sample and N_p the total number of pairs in the sample. Then [35]

$$\gamma_{\mathbf{x},\mathbf{y}} = \frac{N_p^{+} - N_p^{-}}{N_p^{+} + N_p^{-}} = \frac{N_p^{+} - N_p^{-}}{N_p} \cdot \frac{N_p}{N_p^{+} + N_p^{-}} = \tau_{\mathbf{x},\mathbf{y}} \div \frac{N_p^{+} + N_p^{-}}{N_p}. \tag{2.6}$$

Assume that the sample is defined through two dependent random variables X and Y. Let (X_1, Y_1) and (X_2, Y_2) be two independent copies of (X, Y). Following the definitions given by 2.4 and 2.5, denote event $C = A \vee B$ and define

$$p_3 := \mathrm{P}(C). \tag{2.7}$$

Goodman and Kruskal's Gamma for random variables X and Y can be defined as a difference between two conditional probabilities:

$$\gamma_{X,Y} = \mathrm{P}(A \mid C) - \mathrm{P}(B \mid C).$$

Observe that

$$\mathrm{P}(A \mid C) = \frac{\mathrm{P}(A \cap C)}{\mathrm{P}(C)} = \frac{\mathrm{P}(A)}{\mathrm{P}(C)},$$

and

$$\mathrm{P}(B \mid C) = \frac{\mathrm{P}(B \cap C)}{\mathrm{P}(C)} = \frac{\mathrm{P}(B)}{\mathrm{P}(C)},$$

which means that $\gamma_{X,Y}$ can be expressed as

$$\gamma_{X,Y} = \frac{p_1 - p_2}{p_3}.$$

This is in line with the interpretation of γ as a ratio between τ and the proportion of pairwise comparable observations in the sample (highlighted by Equality 2.6).

Corollary 2.14. *Let (X_1, Y_1) and (X_2, Y_2) be two independent copies of (X, Y), where X and Y are dependent random variables. Then*

$$\gamma_{X,Y} = \frac{\mathbb{E}[\mathrm{sgn}(X_1 - X_2) \cdot \mathrm{sgn}(Y_1 - Y_2)]}{\mathrm{P}(X_1 \neq X_2 \wedge Y_1 \neq Y_2)}.$$

2.3 Somers' D

2.3.1 Somers' D for given samples

It is evident with $\tau_{\mathbf{x},\mathbf{y}}$ and $\gamma_{\mathbf{x},\mathbf{y}}$ that the concentration of tied pairs in the sample (across either of the two variables) is not directly accounted for and those quantities rather represent the borderline cases of counting the maximum possible score. It turns out that in case of non-trivial samples, $\tau_{\mathbf{x},\mathbf{x}}$ could be treated as a denominator to adjust the Kendall's rank correlation coefficient. The quantity $\mathrm{D}(\mathbf{y} \mid \mathbf{x})$ was firstly introduced and examined by Robert H. Somers in his paper [35] that was published in 1962. Such an adjustment can be interpreted as excluding all the pairs where the ordering according to $\mathbf{x}$ cannot be determined from the calculation of the maximum possible score, whilst still penalizing for cases where the ordering according to $\mathbf{y}$ cannot be determined.

Definition 2.15. *Let $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ be a sample of observations, $n \in \mathbb{N}$, $n \geq 2$, $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$. Somers' D of $\mathbf{y}$ with respect to $\mathbf{x}$ is defined as*

$$\mathrm{D}(\mathbf{y} \mid \mathbf{x}) := \frac{\tau_{\mathbf{x},\mathbf{y}}}{\tau_{\mathbf{x},\mathbf{x}}}.$$

Whilst both γ and τ were symmetric quantities, i.e. $\gamma_{\mathbf{x},\mathbf{y}} = \gamma_{\mathbf{y},\mathbf{x}}$ and $\tau_{\mathbf{x},\mathbf{y}} = \tau_{\mathbf{y},\mathbf{x}}$ for any given $(\mathbf{x}, \mathbf{y})$, then Somers' D is not. This means that $\mathrm{D}(\mathbf{y} \mid \mathbf{x}) \neq \mathrm{D}(\mathbf{x} \mid \mathbf{y})$ in most practical situations. Whichever of the two quantities is higher can determined by comparing the proportion of pairwise tied observations across vectors $\mathbf{x}$ and $\mathbf{y}$ (or across the columns and rows of a given contingency table). If there are no ties, i.e. $x_i \neq x_j$ and $y_i \neq y_j$ $\forall i, j \in \{1, ..., n\}$ $(i \neq j)$, then $\tau_{\mathbf{x},\mathbf{y}} = \mathrm{D}(\mathbf{y} \mid \mathbf{x}) = \mathrm{D}(\mathbf{x} \mid \mathbf{y}) = \gamma_{\mathbf{x},\mathbf{y}}$. If there are more tied pairs across vector $\mathbf{x}$ compared to $\mathbf{y}$, then $\tau_{\mathbf{x},\mathbf{x}} < \tau_{\mathbf{y},\mathbf{y}}$ and $|\mathrm{D}(\mathbf{y} \mid \mathbf{x})| > |\mathrm{D}(\mathbf{x} \mid \mathbf{y})|$. Similarly, if there are more tied pairs across vector $\mathbf{y}$ compared to $\mathbf{x}$, then $|\mathrm{D}(\mathbf{y} \mid \mathbf{x})| < |\mathrm{D}(\mathbf{x} \mid \mathbf{y})|$.

Corollary 2.16. *Let* $\mathbf{x} = (x_1, ..., x_n)$ *be a sample of observations,* $n \in \mathbb{N}$, $n \geq 2$, $x_i \in \mathbb{R}$ $\forall i \in \{1, ..., n\}$. *Then* $\mathrm{D}(\mathbf{x} \mid \mathbf{x}) = 1$.

The results presented by the following Lemma are not new results discovered by the author. However, as it was not simple to find examples of formal proofs for these statements based on the mathematical construct used in this thesis, an independent proof is provided by the author.

Lemma 2.17. *Let* $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ *be a sample of observations,* $n \in \mathbb{N}$, $n \geq 2$, $x_i, y_i \in \mathbb{R}$ $\forall i \in \{1, ..., n\}$. *Then* $|\tau_{\mathbf{x},\mathbf{y}}| \leq |\mathrm{D}(\mathbf{y} \mid \mathbf{x})| \leq |\gamma_{\mathbf{x},\mathbf{y}}|$.

Proof.
Firstly, as $\tau_{\mathbf{x},\mathbf{x}} \in (0, 1]$ for non-trivial samples, then the adjustment can only increase the absolute value of the initially obtained correlation coefficient, i.e.

$$|\mathrm{D}(\mathbf{y} \mid \mathbf{x})| = \frac{|\tau_{\mathbf{x},\mathbf{y}}|}{\tau_{\mathbf{x},\mathbf{x}}} \geq |\tau_{\mathbf{x},\mathbf{y}}|.$$

Secondly, from

$$\sum_{i=1}^{n-1} \sum_{j>i} [\mathrm{sgn}(x_j - x_i)]^2 \geq \sum_{i=1}^{n-1} \sum_{j>i} [\mathrm{sgn}(x_j - x_i)]^2 \cdot [\mathrm{sgn}(y_j - y_i)]^2,$$

it follows that

$$|\gamma_{\mathbf{x},\mathbf{y}}| = \frac{|\sum_{i=1}^{n-1} \sum_{j>i} \mathrm{sgn}(x_j - x_i) \cdot \mathrm{sgn}(y_j - y_i)|}{\sum_{i=1}^{n-1} \sum_{j>i} [\mathrm{sgn}(x_j - x_i)]^2 \cdot [\mathrm{sgn}(y_j - y_i)]^2} \geq \frac{|\sum_{i=1}^{n-1} \sum_{j>i} \mathrm{sgn}(x_j - x_i) \cdot \mathrm{sgn}(y_j - y_i)|}{\sum_{i=1}^{n-1} \sum_{j>i} [\mathrm{sgn}(x_j - x_i)]^2} = \frac{|\tau_{\mathbf{x},\mathbf{y}}|}{\tau_{\mathbf{x},\mathbf{x}}}.$$

□

Corollary 2.18. *Let* $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ *be a sample of observations,* $n \in \mathbb{N}$, $n \geq 2$, $x_i, y_i \in \mathbb{R}$ $\forall i \in \{1, ..., n\}$. *Then* $|\tau_{\mathbf{x},\mathbf{y}}| \leq |\mathrm{D}(\mathbf{x} \mid \mathbf{y})| \leq |\gamma_{\mathbf{x},\mathbf{y}}|$.

As it was with previously introduced quantities, Somers' D can also be calculated through a contingency table, which is a direct requirement of [21]. Denote the sum of observations in row i as

$$R_i = \sum_{j=1}^{M_2} N_{ij}.$$

Observe that

$$\frac{\sum_{j=1}^{M_2} N_{ij} \cdot (\sum_{j=1}^{M_2} N_{ij} - 1)}{2} = \frac{(\sum_{j=1}^{M_2} N_{ij})^2 - \sum_{j=1}^{M_2} N_{ij}}{2},$$

represents the total number of pairs in the sample that are situated in row i of the contingency table. Summarizing across all the rows of the contingency table yields

$$\frac{\sum_{i=1}^{M_1} \left[(\sum_{j=1}^{M_2} N_{ij})^2 - \sum_{j=1}^{M_2} N_{ij} \right]}{2},$$

which gives us the number of pairs in the sample where ordering according to vector $\mathbf{x}$ cannot be determined. Subtracting this from the total number of pairs in the sample leads to

$$\frac{n \cdot (n-1) - \sum_{i=1}^{M_1} \left[(\sum_{j=1}^{M_2} N_{ij})^2 - \sum_{j=1}^{M_2} N_{ij} \right]}{2} = \frac{n \cdot (n-1) - \sum_{i=1}^{M_1} (\sum_{j=1}^{M_2} N_{ij})^2 + n}{2} =$$

$$= \frac{(\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij})^2 - \sum_{i=1}^{M_1} R_i^2}{2},$$

which represents the number of pairs in the sample if pairs belonging to the same row of the contingency table are discarded. Thus,

$$\mathrm{D}(\mathbf{y} \mid \mathbf{x}) = \frac{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij} \cdot N_{ij}^{+} - \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij} \cdot N_{ij}^{-}}{(\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij})^2 - \sum_{i=1}^{M_1} R_i^2}. \tag{2.8}$$

Similarly, if

$$C_j = \sum_{i=1}^{M_1} N_{ij},$$

represents the sum of observations in column j, then $\mathrm{D}(\mathbf{x} \mid \mathbf{y})$ can be expressed as

$$\mathrm{D}(\mathbf{x} \mid \mathbf{y}) = \frac{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij} \cdot N_{ij}^{+} - \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij} \cdot N_{ij}^{-}}{(\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij})^2 - \sum_{j=1}^{M_2} C_j^2}.$$

When comparing the denominator presented in Equality 2.8 to Equality 2.2, the relation between Somers' D and τ becomes evident by observing that

$$n \cdot (n-1) = (\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij})^2 - \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij} \geq (\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij})^2 - \sum_{i=1}^{M_1} (R_i)^2.$$

Example 2.19. *Let* $\mathbf{x} = (2, 2, 3, 3)$ *and* $\mathbf{y} = (1, 2, 1, 3)$. *Then* $\tau_{\mathbf{x},\mathbf{x}} = \frac{4}{6} = \frac{2}{3}$ *and* $\tau_{\mathbf{y},\mathbf{y}} = \frac{5}{6}$. *The fact that* $\tau_{\mathbf{y},\mathbf{y}} > \tau_{\mathbf{x},\mathbf{x}}$ *can be interpreted as* $\mathbf{y}$ *having more pairwise comparable elements than* $\mathbf{x}$*, which also means that* $\mid \mathrm{D}(\mathbf{y} \mid \mathbf{x}) \mid > \mid \mathrm{D}(\mathbf{x} \mid \mathbf{y}) \mid$.

2.3.2 Somers' D for random samples

Let $N_p^{\mathbf{x}}$ denote the number of pairs in the sample that are tied on $\mathbf{x}$ and let $N_p^{\mathbf{y}}$ denote the number of pairs in the sample that are tied on $\mathbf{y}$. Then [35]

$$\mathrm{D}(\mathbf{y} \mid \mathbf{x}) = \frac{N_p^+ - N_p^-}{N_p - N_p^{\mathbf{x}}} = \frac{N_p^+ - N_p^-}{N_p} \cdot \frac{N_p}{N_p - N_p^{\mathbf{x}}} = \tau_{\mathbf{x},\mathbf{y}} \div \frac{N_p - N_p^{\mathbf{x}}}{N_p}, \tag{2.9}$$

and

$$\mathrm{D}(\mathbf{x} \mid \mathbf{y}) = \frac{N_p^+ - N_p^-}{N_p - N_p^{\mathbf{y}}} = \frac{N_p^+ - N_p^-}{N_p} \cdot \frac{N_p}{N_p - N_p^{\mathbf{y}}} = \tau_{\mathbf{x},\mathbf{y}} \div \frac{N_p - N_p^{\mathbf{y}}}{N_p}, \tag{2.10}$$

where N_p^+ and N_p^- denote the number of concordant and disconcordant pairs in the sample, respectively, and N_p denotes the total number of pairs in the sample. Assume that the sample is defined through two dependent random variables X and Y. Let (X_1, Y_1) and (X_2, Y_2) be two independent copies of (X, Y). Based on the interpretation of Somers' D as a ratio between τ and the proportion of pairwise comparable observations according to $\mathbf{x}$ (highlighted by Equation 2.9) or $\mathbf{y}$ (highlighted by Equality 2.10) and following the definitions given by 2.4 and 2.5, define the *Somers'* D *of random variable Y with respect to random variable X* as

$$\mathrm{D}(Y \mid X) = \frac{p_1 - p_2}{\mathrm{P}(X_1 \neq X_2)}.$$

and similarly, the *Somers'* D *of random variable X with respect to random variable Y* as

$$\mathrm{D}(X \mid Y) = \frac{p_1 - p_2}{\mathrm{P}(Y_1 \neq Y_2)}.$$

Corollary 2.20. *Let (X_1, Y_1) and (X_2, Y_2) be two independent copies of (X, Y), where X and Y are dependent random variables. Then*

$$\mathrm{D}(Y \mid X) = \frac{\mathbb{E}[\mathrm{sgn}(X_1 - X_2) \cdot \mathrm{sgn}(Y_1 - Y_2)]}{\mathrm{P}(X_1 \neq X_2)},$$

and

$$\mathrm{D}(X \mid Y) = \frac{\mathbb{E}[\mathrm{sgn}(X_1 - X_2) \cdot \mathrm{sgn}(Y_1 - Y_2)]}{\mathrm{P}(Y_1 \neq Y_2)}.$$

2.4 Generalized AUC

The current Section highlights an important area of application for Somers' D that is related to the validation of IRB LGD models discussed in Section 1.2.1. We examine the measure gAUC which is defined in the regulatory reporting templates [21] and is based on a linear transformation of Somers' D. Additionally, a general overview of the requirements for reporting the results of discriminatory power assessment of LGD models is provided.

Given a back-testing sample from the relevant observation period, denote the vector of discretized LGD estimates by $\mathbf{LGD^E}$ and the vector of discretized realized LGDs by $\mathbf{LGD^R}$. The discretization of LGDs is performed on facility grade or pool level, depending on the institution's LGD model. An illustration of the discretization process for continuous LGD models according to [21] is given in Chapter 3. The *generalized AUC for sample* $(\mathbf{LGD^E}, \mathbf{LGD^R})$ is defined as

$$gAUC = \frac{\mathrm{D}(\mathbf{LGD^R} \mid \mathbf{LGD^E}) + 1}{2}.$$

It is clear that $1 \geq gAUC \geq 0$ for any non-trivial sample. *gAUC* values close to 1 or 0 indicate a strong correlation as they reflect situations where Somers' D is close to its maximum or minimum value respectively. $gAUC = \frac{1}{2}$ represents the situation when Somers' D is equal to 0 and no correlation between the input parameters was detected.

The name gAUC indicates that the measure relates to the receiver operating characteristic (ROC) curve, as the area under the ROC curve is often referred to as AUC (area under curve). gAUC generalizes the AUC in a sense that AUC deals with binary contingency tables whilst gAUC deals with $M_1 \times M_2$ contingency tables. However, the measures are not directly related due to the fact that in the calculation of *gAUC*, tied pairs across one input vector are ignored. Moreover, the generalization for $M_1 \times M_2$ contingency tables means that it is not possible to visualize the ROC curve in set-ups where M_1 or M_2 are higher than 2. A common way to illustrate the result of *gAUC* is a visualization of the underlying contingency table. An alternative of the traditional AUC in terms of visualization of the proportion of loss in the portfolio that is captured by LGDs could be the *Loss Capture Ratio* (LCR), which is similar to the Cumulative Accuracy Profile (CAP) and explained in detail in the article "Validation techniques and performance metrics for loss given default models" by Li et al. [31].

The main validation tool that is used on an annual basis is the test statistic S, which is defined as

$$S = \frac{gAUC_{\text{init}} - gAUC_{\text{curr}}}{s},$$

where $gAUC_{\text{init}}$ represents the generalized AUC at the moment of initial model validation, $gAUC_{\text{curr}}$ the generalized AUC at the moment of reporting and s the standard deviation of the latter, which can be estimated as

$$s = \frac{1}{w_{\text{row}}^2} \sqrt{\sum_{i=1}^{M_*} \sum_{j=1}^{M_*} N_{ij} \cdot [w_{\text{row}} \cdot d_{ij} - (P - Q) \cdot (F - R_i)]^2},$$

where

$$M_* = \max(M_1, M_2),$$

$$w_{\text{row}} = F^2 - \sum_{i=1}^{M_*} R_i^2,$$

$$d_{ij} = N_{ij}^+ - N_{ij}^-,$$

$$P = \sum_{i=1}^{M_*} \sum_{j=1}^{M_*} N_{ij} \cdot N_{ij}^+,$$

$$Q = \sum_{i=1}^{M_*} \sum_{j=1}^{M_*} N_{ij} \cdot N_{ij}^-,$$

and

$$F = \sum_{i=1}^{M_*} \sum_{j=1}^{M_*} N_{ij}.$$

Each institution is to report on an annual basis the values of both $gAUC_{\text{init}}$ and $gAUC_{\text{curr}}$, the estimated variance s^2, the test statistic S and the p-value $1 - \phi(S)$. In addition, institutions are required to provide information on the used sample (e.g. the observation horizon and variance within the current sample) and a contingency table depicting model estimates and realizations across predefined LGD pools or segments. The contingency table that is provided can be directly used to estimate $gAUC_{\text{curr}}$ and is filled treating the discretized realized losses of the sample as the dependent variable (i.e. as the column-variable) and the discretized model estimates as the independent variable (i.e. as the row-variable).

The test statistic S can be used to assess whether the discriminatory power of the model has deteriorated compared to initial validation. In particular, the null hypothesis for the test is that $gAUC_{\text{init}}$ is lower than $gAUC_{\text{curr}}$. However, the reporting instructions do not define any thresholds for assessing the results of the underlying $gAUC$ measures standalone. Although not explicitly required by the regulation, it makes sense from a practical perspective for financial institutions with IRB permission to incorporate the test into their annual validation methodology, as those results have to be reported each year together with the validation results. This means that an assessment scale has to be developed for the interpretation of the test results.

For linear models the square of the Pearson's correlation coefficient (see Section 2.6.3 for the definition) may be interpreted as the proportion of variance in one variable that is accounted for by the differences in the other variable [34]. This relationship can be used as a basis to set the borders of acceptable Pearson's correlation for linear models. It is also noted by [34] that the square of Pearson's correlation coefficient can be expressed as a product of the slopes of the linear regression lines for predicting one variable from the other. However, even though a similar relationship (as between Pearson's correlation and the slopes of the regression lines) holds between τ^b defined in Section 2.6.1 and Somers' D (see Corollary 2.28), such interpretation is not as useful when dealing with monotonic correlations [35]. In general, the interpretation of gAUC depends on the circumstances (sample and / or objective) and the final decision in terms of model acceptance could be supported with additional investigation. This includes (but is not limited to) the analysis across alternative measures of discriminatory power, extension of the observation-horizon, qualitative examination of the data sample, expert-based review of the current and previous versions of the model and benchmarking against competitors or challenger models.

Section 2.5 below illustrates one particular example where borderline results can be observed in practice: the case of low-loss portfolios.

2.5 Samples with significant concentrations of tied pairs

The fact that realized LGD is treated as a dependent variable ($\mathbf{y}$) for the calculation of $\mathrm{D}(\mathbf{y} \mid \mathbf{x})$ in the regulatory reporting templates ([21]) means that all the comparisons of observations where the realized losses belong to the same bucket but estimated LGD buckets differ, increase the denominator in the calculation yet leave the numerator unchanged. In the opposite situation where the estimated LGDs belong to the same bucket, but realized loss buckets differ, neither the denominator nor the numerator is increased. This leads to an understanding that in certain situations, such as for low-loss portfolios, the measure proposed by the regulatory reporting templates might not be able to reach its maximum value (1) even if all the pairwise comparable elements in the sample are concordant.

Example 2.21. *Assume that the discretized LGD estimates and realized losses of 20 defaults are depicted in the following contingency table, where variable x represents the model estimates and variable y the realized losses. For simplicity, the discretization could be thought of as dividing the estimates and realizations across three equal intervals:* $[0, \frac{1}{3})$, $[\frac{1}{3}, \frac{2}{3})$ *and* $[\frac{2}{3}, 1]$.

x \ y	1	2	3
1	9	0	0
2	5	1	1
3	2	1	1

Table 2.2: The contingency table for Example 2.21

For the given sample, $\tau_{\mathbf{x},\mathbf{x}} \approx 0.67$, $\tau_{\mathbf{y},\mathbf{y}} \approx 0.36$, $\tau_{\mathbf{x},\mathbf{y}} \approx 0.22$, $\gamma_{\mathbf{x},\mathbf{y}} \approx 0.81$, $\mathrm{D}(\mathbf{y} \mid \mathbf{x}) \approx 0.33$ *and* $\mathrm{D}(\mathbf{x} \mid \mathbf{y}) \approx 0.62$. *The fact that* $\tau_{\mathbf{x},\mathbf{x}} > \tau_{\mathbf{y},\mathbf{y}}$ *means that there are more pairwise-tied observations across vector* $\mathbf{y}$ *than across vector* $\mathbf{x}$*. In fact, as* $\tau_{\mathbf{y},\mathbf{y}} < 0.5$*, then more than half of the comparisons performed on* $\mathbf{y}$ *are tied. The fact that the sample contains a high concentration of tied pairs is also illustrated by the noticeable difference between the values of* $\tau_{\mathbf{x},\mathbf{y}}$ *and* $\gamma_{\mathbf{x},\mathbf{y}}$*. The relatively high value of* $\gamma_{\mathbf{x},\mathbf{y}}$ *indicates that most of the non-tied pairs in the sample are concordant.*

When comparing the quantities that were presented in the example above it is essential to note that standard deviations of the measures for the sample vary as the number of comparisons performed (i.e. the maximum possible scores when interpreted similarly to $\tau_{\mathbf{x}}$) differ. For example, when calculating $\mathrm{D}(\mathbf{y} \mid \mathbf{x})$, the number of performed comparisons is significantly higher compared to $\mathrm{D}(\mathbf{x} \mid \mathbf{y})$, which is indicated by the fact that $\tau_{\mathbf{x},\mathbf{x}} > \tau_{\mathbf{y},\mathbf{y}}$. The relatively high value of $\mathrm{D}(\mathbf{x} \mid \mathbf{y})$ should be interpreted with caution, taking into consideration that the number of comparisons made in practice when estimating the measure is low.

One way to illustrate the bias towards τ that stems from having tied pairs in the sample is to calculate the *maximum value of* τ given the concentration of tied pairs in the sample (τ^{max}).

Definition 2.22. *Let* $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ *be a sample of observations,* $n \in \mathbb{N}$, $n \geq 2$, $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$*. The maximum value of* τ *for sample* $(\mathbf{x}, \mathbf{y})$*, given the concentration of tied pairs in the sample is defined as*

$$\tau^{max}_{\mathbf{x},\mathbf{y}} = \frac{2}{(n-1) \cdot n} \cdot \sum_{i=1}^{n-1} \sum_{j>i} [\mathrm{sgn}(x_j - x_i)]^2 \cdot [\mathrm{sgn}(y_j - y_i)]^2.$$

$\tau^{max}_{\mathbf{x},\mathbf{y}}$ can also be directly calculated from contingency tables following the definitions presented in Section 2.1.1.2:

$$\tau^{max}_{\mathbf{x},\mathbf{y}} = \frac{\sum_{i=1}^{M_1}\sum_{j=1}^{M_2} N_{ij}\cdot(N^+_{ij}+N^-_{ij})}{(n-1)\cdot n}.$$

The aim of calculating τ^{max} is not to provide an alternative correlation coefficient to τ, but to understand the impact from having tied pairs in the sample. The closer the value of τ^{max} is to 0, the more significant the concentration of tied pairs.

Corollary 2.23. *Let* $(\mathbf{x},\mathbf{y}) = ((x_1,y_1),...,(x_n,y_n))$ *be a sample of observations,* $n\in\mathbb{N}$, $n\geq 2$, $x_i, y_i \in \mathbb{R}\ \forall i\in\{1,...,n\}$. *Then* $\tau^{max}_{\mathbf{x},\mathbf{y}} = \frac{\tau_{\mathbf{x},\mathbf{y}}}{\gamma_{\mathbf{x},\mathbf{y}}}$.

The definition of $\tau^{max}_{\mathbf{x},\mathbf{y}}$ can be extended for Somers' D in the following way.

Definition 2.24. *Let* $(\mathbf{x},\mathbf{y}) = ((x_1,y_1),...,(x_n,y_n))$ *be a sample of observations,* $n\in\mathbb{N}$, $n\geq 2$, $x_i, y_i \in \mathbb{R}\ \forall i\in\{1,...,n\}$. *The maximum value of* $\mathrm{D}(\mathbf{y}\mid\mathbf{x})$ *for sample* $(\mathbf{x},\mathbf{y})$, *given the concentration of tied pairs in the sample is defined as*

$$D^{max}_{\mathbf{y}|\mathbf{x}} = \frac{\tau^{max}_{\mathbf{x},\mathbf{y}}}{\tau_{\mathbf{x},\mathbf{x}}}.$$

Example 2.25. *Let* $(\mathbf{x},\mathbf{y}) = ((2,1),(2,2),(3,1),(3,3))$. *In order to estimate* $\tau^{max}_{\mathbf{x},\mathbf{y}}$, *the sum of both concordant and disconcordant observations has to be calculated. According to 2.23, 2.6 and 2.13* $\tau^{max}_{\mathbf{x},\mathbf{y}} = \frac{1}{2}$. *This result indicates that half of the comparisons performed involve tied pairs. As* $\tau_{\mathbf{x},\mathbf{x}} = \frac{2}{3}$, *then* $D^{max}_{\mathbf{y}|\mathbf{x}} = \frac{3}{4}$.

As it was with τ^{max}, D^{max} should not be seen as an alternative correlation coefficient to Somers' D, but as an indicator to understand the impact from having tied pairs in the sample. However, if $D^{max}_{\mathbf{y}|\mathbf{x}}$ is treated as an adjustment to the originally obtained correlation coefficient $\mathrm{D}(\mathbf{y}\mid\mathbf{x})$, then the adjusted correlation coefficient is equal to Goodman and Kruskal's gamma:

$$\frac{\mathrm{D}(\mathbf{y}\mid\mathbf{x})}{D^{max}_{\mathbf{y}|\mathbf{x}}} = \frac{\tau_{\mathbf{x},\mathbf{y}}}{\tau_{\mathbf{x},\mathbf{x}}}\cdot\frac{\tau_{\mathbf{x},\mathbf{x}}}{\tau^{max}_{\mathbf{x},\mathbf{y}}} = \gamma_{\mathbf{x},\mathbf{y}}.$$

The result above indicates that for samples with significant concentration of tied pairs (such as low-loss portfolios), γ could be used as a supportive measure when assessing the results of *gAUC* during model development / validation. For instance, even though $\mathrm{D}(\mathbf{y}\mid\mathbf{x}) \approx 0.33$ in Example 2.21 indicates a rather mild correlation, the fact that $\gamma_{\mathbf{x},\mathbf{y}} \approx 0.81$ for the same sample means that among pairwise-comparable observations in the sample, the correlation is quite strong. If the volume of pairwise-comparable observations in the sample is deemed sufficient and the model performs adequately in all other aspects, the relatively low value of $\mathrm{D}(\mathbf{y}\mid\mathbf{x})$ should not dismiss the developed model in such cases.

2.6 Additional measures of correlation for LGD models

In this section, three additional measures of correlation are introduced, which will be used in comparison to the previously defined correlation coefficients in the simulation presented in

Chapter 3.

In his book "Rank Correlation Methods" [30], Maurice George Kendall discusses an adjustment to the initially developed correlation coefficient τ. The adjustment is applied on the denominator of the correlation coefficient: instead of accounting for all the pairs in the calculation of the maximum possible score, the geometric mean of pairwise comparable observations according to vectors $\mathbf{x}$ and $\mathbf{y}$ could be considered. The resulting correlation coefficient is referred to as τ^b.

In the same book, a second adjustment for the denominator of τ is examined. The second adjustment is based on the number of unique observations across vectors $\mathbf{x}$ and $\mathbf{y}$, whichever is lower. The resulting correlation coefficient is referred to as τ^c.

Lastly, we define a commonly used measure of association, which is not directly related to τ: the Spearman's rank-correlation coefficient [13].

2.6.1 τ^b for given samples

Let N_p^+ denote the number of concordant pairs in the sample, N_p^- the number of disconcordant pairs in the sample and N_p the total number of pairs in the sample. Denote the number of pairs in the sample that are tied on x by $N_p^{\mathbf{x}}$ and the number of pairs in the sample that are tied on y by $N_p^{\mathbf{y}}$.

Definition 2.26. *Let $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ be a sample of observations, $n \in \mathbb{N}$, $n \geq 2$, $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$. Then τ^b for sample $(\mathbf{x}, \mathbf{y})$ is defined as*

$$\tau^b_{\mathbf{x},\mathbf{y}} = \frac{N_p^+ - N_p^-}{\sqrt{(N_p - N_p^{\mathbf{x}}) \cdot (N_p - N_p^{\mathbf{y}})}}. \tag{2.11}$$

A comparison of denominators of Equalities 2.11, 2.6 and 2.3 provides a basic understanding of how τ^b relates to τ and γ. It must be noted that the result presented by the following two Lemmas are not new results discovered by the author. As it was not simple to find examples of formal proofs for these statements based on the mathematical construct used in this thesis, independent proofs were constructed.

Lemma 2.27. *Let $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ be a sample of observations, $n \in \mathbb{N}$, $n \geq 2$, $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$. Then $|\tau_{\mathbf{x},\mathbf{y}}| \leq |\tau^b_{\mathbf{x},\mathbf{y}}| \leq |\gamma_{\mathbf{x},\mathbf{y}}|$.*

Proof.
Firstly, observe that

$$\sqrt{(N_p - N_p^{\mathbf{x}}) \cdot (N_p - N_p^{\mathbf{y}})} \leq \sqrt{N_p \cdot N_p} = N_p,$$

which means that $|\tau^b_{\mathbf{x},\mathbf{y}}| \geq |\tau_{\mathbf{x},\mathbf{y}}|$ as the enumerators for both correlation coefficients are equal (see Equalities 2.11 and 2.3).

Next, denote the number of pairs in the sample that are tied on both x and y by $N_p^{\mathbf{x},\mathbf{y}}$ and observe that the square of the denominator in Equality 2.11 can be expressed as

$$(N_p - N_p^{\mathbf{x}}) \cdot (N_p - N_p^{\mathbf{y}}) = (N_p^+ + N_p^- + N_p^{\mathbf{y}} - N_p^{\mathbf{x},\mathbf{y}}) \cdot (N_p^+ + N_p^- + N_p^{\mathbf{x}} - N_p^{\mathbf{x},\mathbf{y}}).$$

Due to the fact that $N_p^{\mathbf{y}} - N_p^{\mathbf{x},\mathbf{y}} \geq 0$ and $N_p^{\mathbf{x}} - N_p^{\mathbf{x},\mathbf{y}} \geq 0$, it is evident that

$$(N_p - N_p^{\mathbf{x}}) \cdot (N_p - N_p^{\mathbf{y}}) \geq (N_p^+ + N_p^-) \cdot (N_p^+ + N_p^-),$$

which means that the denominator of γ is less than or equal to the denominator of τ^b (see Equalities 2.11 and 2.6). Hence, $|\gamma_{\mathbf{x},\mathbf{y}}| \geq |\tau^b_{\mathbf{x},\mathbf{y}}|$. □

It was noted by Robert H. Somers in his paper [35] that the following relation to $\mathrm{D}(\mathbf{y} \mid \mathbf{x})$ and $\mathrm{D}(\mathbf{x} \mid \mathbf{y})$ holds, which is evident by comparing Equalities 2.9 and 2.10 to Equality 2.11.

Corollary 2.28. *Let* $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ *be a sample of observations,* $n \in \mathbb{N}$, $n \geq 2$, $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$. *Then* $(\tau^b_{\mathbf{x},\mathbf{y}})^2 = \mathrm{D}(\mathbf{y} \mid \mathbf{x}) \cdot \mathrm{D}(\mathbf{x} \mid \mathbf{y})$.

The direct relation between $\tau^b_{\mathbf{x},\mathbf{y}}$, $\mathrm{D}(\mathbf{y} \mid \mathbf{x})$ and $\mathrm{D}(\mathbf{x} \mid \mathbf{y})$ depends on the proportion of tied pairs across vectors $\mathbf{x}$ and $\mathbf{y}$.

Lemma 2.29. *Let* $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ *be a sample of observations,* $n \in \mathbb{N}$, $n \geq 2$, $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$. *Then* $|\mathrm{D}(\mathbf{y} \mid \mathbf{x})| \leq |\tau^b_{\mathbf{x},\mathbf{y}}| \leq |\mathrm{D}(\mathbf{x} \mid \mathbf{y})|$ *or* $|\mathrm{D}(\mathbf{x} \mid \mathbf{y})| \leq |\tau^b_{\mathbf{x},\mathbf{y}}| \leq |\mathrm{D}(\mathbf{y} \mid \mathbf{x})|$.

Proof.
If the number of tied pairs is equal across vectors $\mathbf{x}$ and $\mathbf{y}$, i.e. $N_p^{\mathbf{x}} = N_p^{\mathbf{y}}$, then the equality $\mathrm{D}(\mathbf{y} \mid \mathbf{x}) = \mathrm{D}(\mathbf{x} \mid \mathbf{y}) = \tau^b_{\mathbf{x},\mathbf{y}}$ holds. If there are more tied pairs across vector $\mathbf{y}$ compared to $\mathbf{x}$, then $N_p - N_p^{\mathbf{y}} < N_p - N_p^{\mathbf{x}}$ and

$$\sqrt{(N_p - N_p^{\mathbf{y}})^2} < \sqrt{(N_p - N_p^{\mathbf{x}}) \cdot (N_p - N_p^{\mathbf{y}})} < \sqrt{(N_p - N_p^{\mathbf{x}})^2},$$

which means that $|\mathrm{D}(\mathbf{y} \mid \mathbf{x})| < |\tau^b_{\mathbf{x},\mathbf{y}}| < |\mathrm{D}(\mathbf{x} \mid \mathbf{y})|$ (see Equalities 2.9, 2.10 and 2.11). Similarly, if there are more tied pairs across vector $\mathbf{x}$ compared to $\mathbf{y}$, then $|\mathrm{D}(\mathbf{y} \mid \mathbf{x})| > |\tau^b_{\mathbf{x},\mathbf{y}}| > |\mathrm{D}(\mathbf{x} \mid \mathbf{y})|$. □

The quantity $\tau^b_{\mathbf{x},\mathbf{y}}$ can be expressed through contingency tables [10], following the definitions presented in Sections 2.1.1.2 and 2.3.1 (see Equality 2.8):

$$\tau^b_{\mathbf{x},\mathbf{y}} = \frac{\sum_{i=1}^{M_1}\sum_{j=1}^{M_2} N_{ij} \cdot N_{ij}^+ - \sum_{i=1}^{M_1}\sum_{j=1}^{M_2} N_{ij} \cdot N_{ij}^-}{\sqrt{\left[(\sum_{i=1}^{M_1}\sum_{j=1}^{M_2} N_{ij})^2 - \sum_{i=1}^{M_1} R_i^2\right] \cdot \left[(\sum_{i=1}^{M_1}\sum_{j=1}^{M_2} N_{ij})^2 - \sum_{j=1}^{M_2} C_j^2\right]}}.$$

2.6.2 τ^c for given samples

Definition 2.30. *Let* $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ *be a sample of observations,* $n \in \mathbb{N}$, $n \geq 2$, $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$. *Then* τ^c *for sample* $(\mathbf{x}, \mathbf{y})$ *is defined as*

$$\tau^c_{\mathbf{x},\mathbf{y}} = \frac{2 \cdot m}{n^2 \cdot (m-1)} \cdot \sum_{i=1}^{n-1}\sum_{j>i} \mathrm{sgn}(x_j - x_i) \cdot \mathrm{sgn}(y_j - y_i), \qquad (2.12)$$

where m *represents the number of unique* x_i *or* y_i*, whichever is lower.*

As noted in [30], observe that τ^c can be expressed as

$$\tau^c_{\mathbf{x},\mathbf{y}} = \frac{n-1}{n} \cdot \frac{m}{m-1} \cdot \tau_{\mathbf{x},\mathbf{y}}. \tag{2.13}$$

Due to the fact that $n \geq m$, it is evident that $\frac{n-1}{n} \cdot \frac{m}{m-1} \geq 1$, which allows us to formulate the following.

Corollary 2.31. *Let* $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ *be a sample of observations,* $n \in \mathbb{N}$*,* $n \geq 2$*,* $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$*. Then* $|\tau^c_{\mathbf{x},\mathbf{y}}| \geq |\tau_{\mathbf{x},\mathbf{y}}|$*.*

In the following, we are going to provide a mathematical explanation to the multiplier in front of the summation sign in Formula 2.12, which creates an understanding of how τ^c relates to γ and Somers' D. It must be noted that the result presented by the following Theorem is not a new results discovered by the author. However, as it was not simple to find examples of formal proof for the statement, an independent proof is provided by the author.

Theorem 2.32. *Let* $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ *be a sample of observations,* $n \in \mathbb{N}$*,* $n \geq 2$*,* $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$*. Then* $|\tau^c_{\mathbf{x},\mathbf{y}}| \leq |\gamma_{\mathbf{x},\mathbf{y}}|$*.*

Proof.
The proof was constructed based on the observation made in [30] about the multiplier used in the definition of τ^c (see Formula 2.12), namely

$$\left(\frac{2 \cdot m}{n^2 \cdot (m-1)}\right)^{-1} = \left(\frac{n}{m}\right)^2 \cdot \binom{m}{2}.$$

Assume that there are $m_1 = m$ distinct values across vector $\mathbf{x}$, based on which the population is divided into groups and there are m_2 distinct values in the second vector $\mathbf{y}$, where $1 < m_1 \leq m_2 \leq n$. The proof is similar if the situation is switched. In such set-up there are $\binom{m}{2}$ pairs of distinct groups in the sample, where all the pairs of observations within a single group are neither concordant or disconcordant (because the observations within a single group are tied on $\mathbf{x}$) , but each pair from two distinct groups can be concordant, disconcordant or neither of the two. Denote the number of observations in group i by n_i ($i \in \{1, ..., m\}$). If each observation in the sample is only compared to the observations that do not belong to the same group with them, then the total number of unique comparisons performed across the sample is equal to

$$N_p - N_p^{\mathbf{x}} = \frac{1}{2} \cdot \sum_{i=1}^{m} n_i \cdot (n - n_i) = \frac{1}{2} \cdot \left(n^2 - \sum_{i=1}^{m} n_i^2\right),$$

where N_p represents the total number of pairs across the sample (without grouping) and $N_p^{\mathbf{x}}$ the number of pairs in the sample that are tied on $\mathbf{x}$. The multiplication by $\frac{1}{2}$ is needed in the formula above as by summarizing across all the groups of observations, each of the compared pairs is accounted for twice. Next, note that if we consider the constraint

$$\sum_{i=1}^{m} n_i = n,$$

then the function

$$\mathrm{f}(n_1, ..., n_m) := \sum_{i=1}^{m} n_i^2,$$

obtains its minimum value at $n_1 = ... = n_m = \frac{n}{m}$. If this is the case, then we can write that

$$\max(N_p - N_p^{\mathbf{x}}) = \frac{1}{2} \cdot \sum_{i=1}^{m} \frac{n}{m} \cdot (n - \frac{n}{m}) = \frac{n \cdot (n - \frac{n}{m})}{2}.$$

When both sides of the equation above are divided by N_p, a direct link between Somers' D and τ^c can be established (see Equalities 2.9 and 2.13), namely

$$\frac{\max(N_p - N_p^{\mathbf{x}})}{N_p} = \frac{n \cdot (n - \frac{n}{m})}{2 \cdot N_p} = \frac{2 \cdot n \cdot (n - \frac{n}{m})}{2 \cdot n \cdot (n-1)} = \frac{n \cdot (m-1)}{m \cdot (n-1)},$$

which leads to

$$\min(|\mathrm{D}(\mathbf{y} \mid \mathbf{x})|) = |\tau^c_{\mathbf{x},\mathbf{y}}|.$$

Considering Lemma 2.17, it follows that

$$|\gamma_{\mathbf{x},\mathbf{y}}| \geq |\tau^c_{\mathbf{x},\mathbf{y}}|.$$

If the initial assumption is reversed, i.e. $1 < m_2 \leq m_1 \leq n$, then analogously it can be shown that

$$\min(|\mathrm{D}(\mathbf{x} \mid \mathbf{y})|) = |\tau^c_{\mathbf{x},\mathbf{y}}|,$$

which leads to the same conclusion (see Corollary 2.18). □

Corollary 2.33. *Let* $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ *be a sample of observations,* $n \in \mathbb{N}$, $n \geq 2$, $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$. *Then* $|\tau^c_{\mathbf{x},\mathbf{y}}| \leq |\mathrm{D}(\mathbf{y} \mid \mathbf{x})|$ *or* $|\tau^c_{\mathbf{x},\mathbf{y}}| \leq |\mathrm{D}(\mathbf{x} \mid \mathbf{y})|$.

The fact that $|\tau^c| \leq |\gamma|$ is also mentioned in the article [10]. However, the same source states that $|\tau^c| \leq |\tau^b|$, which is not always the case as shown in the Example 2.34 below. The same example also illustrates a situation where $|\mathrm{D}(\mathbf{x} \mid \mathbf{y})| < |\tau^c_{\mathbf{x},\mathbf{y}}| < |\mathrm{D}(\mathbf{y} \mid \mathbf{x})|$ whilst the results presented in Chapter 3 show that $|\tau^c_{\mathbf{x},\mathbf{y}}|$ can also be smaller than both $|\mathrm{D}(\mathbf{y} \mid \mathbf{x})|$ and $|\mathrm{D}(\mathbf{x} \mid \mathbf{y})|$. Before we proceed with Example 2.34, note that $\tau^c_{\mathbf{x},\mathbf{y}}$ can be expressed through contingency tables [10], following the definitions presented in Section 2.1.1.2:

$$\tau^c_{\mathbf{x},\mathbf{y}} = \frac{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij} \cdot N^+_{ij} - \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij} \cdot N^-_{ij}}{(\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} N_{ij})^2} \cdot \frac{M}{M-1},$$

where $M = \min(M_1, M_2)$.

Example 2.34. *Assume that the discretized LGD estimates and realized losses of 5 defaults are depicted in the following contingency table, where variable* x *represents the model estimates and variable* y *the realized losses.*

X \ Y				
	2	0	1	0
	0	0	1	1

Table 2.3: The contingency table for Example 2.34

For the given sample, $m = 2$*,* $N_p = 10$*,* $N_p^+ = 5$*,* $N_p^- = 0$*,* $N_p^{\mathbf{x}} = 4$ *and* $N_p^{\mathbf{y}} = 2$*. This means that* $\mathrm{D}(\mathbf{y} \mid \mathbf{x}) = \frac{5}{6}$, $\mathrm{D}(\mathbf{x} \mid \mathbf{y}) = \frac{5}{8}$, $\tau^b_{\mathbf{x},\mathbf{y}} = \frac{5}{\sqrt{6 \cdot 8}} \approx \frac{5}{6.93}$ *and* $\tau^c_{\mathbf{x},\mathbf{y}} = \frac{5}{6.25}$*, i.e.* $\mathrm{D}(\mathbf{y} \mid \mathbf{x}) > \tau^c_{\mathbf{x},\mathbf{y}} > \tau^b_{\mathbf{x},\mathbf{y}} > \mathrm{D}(\mathbf{x} \mid \mathbf{y})$.

2.6.3 Spearman's rank-correlation coefficient

Spearman's rank-correlation coefficient is a commonly used tool to measure correlations within samples and assess whether the relationship between the input parameters can be explained by a monotonic function. The correlation coefficient is calculated by ranking the observations across two input variables and then estimating the Pearson's correlation coefficient for those rankings. Thus, we firstly introduce the definition of Pearson's correlation coefficient. [34]

Definition 2.35. *Let* $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ *be a sample of observations,* $n \in \mathbb{N}$*,* $n \geq 2$*,* $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$. Pearson's correlation coefficient for sample $(\mathbf{x}, \mathbf{y})$ *is defined as*

$$r_{\mathbf{x},\mathbf{y}} = \frac{\sum\limits_{i=1}^{n} (x_i - \overline{x}) \cdot (y_i - \overline{y})}{\sqrt{\sum\limits_{i=1}^{n} (x_i - \overline{x})^2 \cdot \sum\limits_{i=1}^{n} (y_i - \overline{y})^2}},$$

where $\overline{x}$ *and* $\overline{y}$ *represent the arithmetic averages of vectors* $\mathbf{x}$ *and* $\mathbf{y}$ *across the sample respectively.*

Similarly to τ, γ and Somers' D, Pearson's correlation coefficient obtains values between -1 and 1. Values close to the upper- or lower- limit indicate a strong linear relation between the input variables.

Whilst Pearson's correlation coefficient explicitly looks for linear relations within the data, Spearman's rank correlation coefficient is not limited to only linear relations, but examines the monotonicity of the relation in general. In order to calculate Spearman's rank correlation coefficient for a sample, the observations in the sample have to be ranked according to both input variables. The rank according to a particular input variable can be derived by sorting the sample in ascending order based on the desired variable and enumerating the observations from left to right. For tied observations, the final rank should be set equal to the arithmetic average of ranks that are occupied by the tied observations [13].

Let a_i denote the rank of observation i based on vector $\mathbf{x}$ and let b_i denote the rank of observation i based on vector $\mathbf{y}$. Denote the arithmetic averages of rankings across the sample as $\overline{a}$ and $\overline{b}$.

Definition 2.36. *Let* $(\mathbf{x}, \mathbf{y}) = ((x_1, y_1), ..., (x_n, y_n))$ *be a sample of observations,* $n \in \mathbb{N}$*,* $n \geq 2$*,* $x_i, y_i \in \mathbb{R}\ \forall i \in \{1, ..., n\}$*.* Spearman's rank correlation coefficient for sample $(\mathbf{x}, \mathbf{y})$ *is defined as*

$$\rho_{\mathbf{x},\mathbf{y}} = \frac{\sum\limits_{i=1}^{n}(a_i - \overline{a}) \cdot (b_i - \overline{b})}{\sqrt{\sum\limits_{i=1}^{n}(a_i - \overline{a})^2 \cdot \sum\limits_{i=1}^{n}(b_i - \overline{b})^2}}.$$

Spearman's rank-correlation coefficient can also be expressed using contingency tables [10], following the definitions presented in Sections 2.1.1.2 and 2.3.1:

$$\rho_{\mathbf{x},\mathbf{y}} = \frac{12 \cdot \sum\limits_{i=1}^{M_1}\sum\limits_{j=1}^{M_2} N_{ij} \cdot \mathbf{R}(i) \cdot \mathbf{C}(j)}{\sqrt{\left[(\sum\limits_{i=1}^{M_1}\sum\limits_{j=1}^{M_2} N_{ij})^3 - \sum\limits_{i=1}^{M_1} R_i^3\right] \cdot \left[(\sum\limits_{i=1}^{M_1}\sum\limits_{j=1}^{M_2} N_{ij})^3 - \sum\limits_{j=1}^{M_2} C_j^3\right]}},$$

where

$$\mathbf{R}(i) = \sum_{k<i} R_k + \frac{R_i}{2} - \frac{\sum\limits_{i=1}^{M_1}\sum\limits_{j=1}^{M_2} N_{ij}}{2},$$

and

$$\mathbf{C}(j) = \sum_{l<j} C_l + \frac{R_i}{2} - \frac{\sum\limits_{i=1}^{M_1}\sum\limits_{j=1}^{M_2} N_{ij}}{2}.$$

3 Simulation of an LGD model

The aim of this chapter is to provide an example of how the different measures of discriminatory power that were described in Chapters 2 and 2.6 compare to one-another in a practical situation. For this purpose, a dataset consisting of financial transactions and their losses was extracted from Kaggle's Loan Default Prediction competition [28]. This data was chosen because it consists of a relatively large number of observations and variables (which allows to develop a sample LGD model), and illustrates the case of a low-loss portfolio. All the operations with the data were performed in R [33].

From the source, two datasets were available: the train dataset (`train_v2.csv`) and the test dataset (`test_v2.csv`). As the observed losses are only presented in the train dataset, then this dataset was chosen for the simulation. The train dataset consists of 105471 observations and 770 variables. The observed loss is indicated by variable `loss` and its values range from 0 (indicating that no loss was observed) to 100 (indicating full loss on the exposure).

Due to the fact that no default indicator is present in the dataset, observations which had carried any loss (i.e. where `loss` > 0) were selected as defaults for model development and backtesting purposes (9783 observations in total). In order to perform out-of-sample validation, the population of defaults was split into development dataset (6522 observations) and validation dataset (3261 observations) by taking a random sample. The sampling was done using `set.seed(10)` to make the results easily replicable.

For the development of the sample LGD model, generalized linear models were considered [36]. The aim of generalized linear models is to predict the expected value of the response variable Y conditional on the independent variables $\mathbf{X}$, assuming that the conditional distribution of the response variable belongs to the exponential family of distributions. The expected value of the response variable is connected to the linear combination of the predictors through a link function g() as follows:

$$\mathbb{E}(Y \mid \mathbf{X}) = \mathrm{g}^{-1}(\mathbf{X} \cdot \boldsymbol{\beta}),$$

where the column vector $\boldsymbol{\beta}$ represents a vector of unknown model parameters and $\mathbf{X}$ is a row vector of independent variables. The parameters for the model are derived using maximum likelihood estimation. This is possible because in such setup, the natural parameters of the exponential family (often referred to as the canonical parameters of the model) which are estimated by maximum likelihood can be directly expressed through the linear combination of $\mathbf{X}$ and $\boldsymbol{\beta}$.

As a first step in model development, the probability density plot was generated for the loss variable in the development dataset.

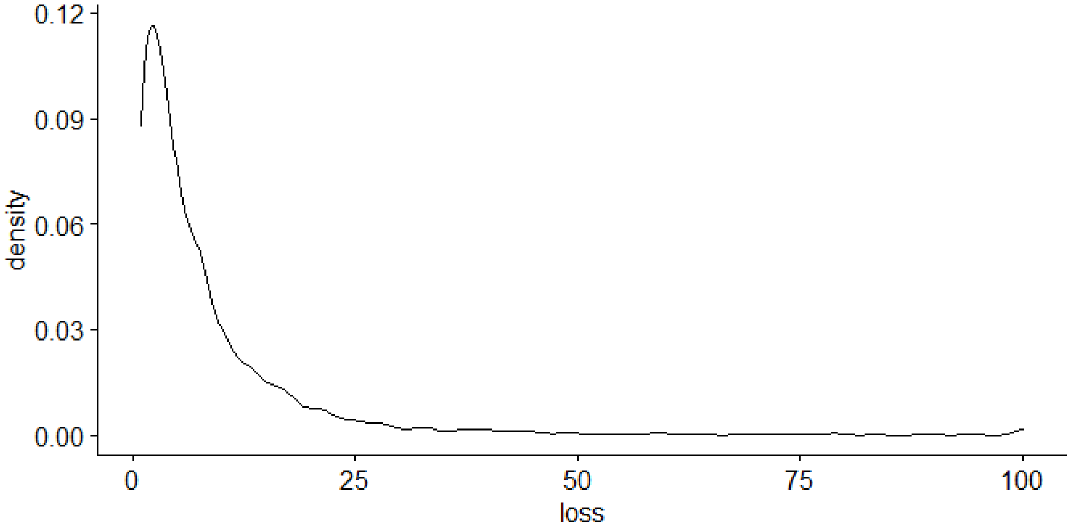

Figure 3.1: The probability density plot for loss on the development dataset.

As can be seen from the figure above, most of the observations in the sample carried a loss below 25 (percent). Visually, the probability density plot indicates that the Gamma distribution, the inverse Gaussian distribution and the lognormal distribution should be considered when fitting generalized linear models.

As a next step, the aim was to narrow the scope of the input variables. Firstly, variables that carried missing values across any of the observations (492 variables in total) were excluded to simplify the model development process. Next, only variables that exhibited the strongest correlation to observed losses were selected. In total, 12 of the remaining input variables yielded a Pearson's correlation coefficient of above 0.3 or below -0.3 towards the loss variable. Those variables were selected for the model development.

	f281	f282	f322	f400	f404	f405	f675	f676	f765	f766	f767	loss
f281	1	0.8874	-0.7576	0.8421	0.8142	0.7291	0.8379	0.6946	0.6946	0.8145	0.7291	0.3664
f282	0.8874	1	-0.6647	0.9946	0.7292	0.8218	0.9883	0.8191	0.8191	0.7296	0.8218	0.332
f322	-0.7576	-0.6647	1	-0.6286	-0.7618	-0.6709	-0.6373	-0.6368	-0.6368	-0.7617	-0.6709	-0.3078
f400	0.8421	0.9946	-0.6286	1	0.6922	0.8165	0.9932	0.8222	0.8222	0.6927	0.8166	0.3159
f404	0.8142	0.7292	-0.7618	0.6922	1	0.8946	0.7008	0.8515	0.8515	0.9996	0.8945	0.3765
f405	0.7291	0.8218	-0.6709	0.8165	0.8946	1	0.8249	0.9948	0.9948	0.8949	1	0.3435
f675	0.8379	0.9883	-0.6373	0.9932	0.7008	0.8249	1	0.8302	0.8302	0.7013	0.8249	0.3217
f676	0.6946	0.8191	-0.6368	0.8222	0.8515	0.9948	0.8302	1	1	0.8519	0.9949	0.3287
f765	0.6946	0.8191	-0.6368	0.8222	0.8515	0.9948	0.8302	1	1	0.8519	0.9949	0.3287
f766	0.8145	0.7296	-0.7617	0.6927	0.9996	0.8949	0.7013	0.8519	0.8519	1	0.8948	0.3761
f767	0.7291	0.8218	-0.6709	0.8166	0.8945	1	0.8249	0.9949	0.9949	0.8948	1	0.3433
loss	0.3664	0.332	-0.3078	0.3159	0.3765	0.3435	0.3217	0.3287	0.3287	0.3761	0.3433	1

Table 3.1: Pearson's correlations between loss and the selected input variables

Variables `f404` and `f766` exhibited the strongest absolute Pearson's correlation with the observed losses (approximately 0.3765 and 0.3761 respectively). As the two variables have a very strong correlation with each other (approximately 0.999), then only variable `f404` was selected for the model. Among the remaining 10 variables, variable `f400` had the absolute correlation below 0.75 with three variables: `f282`, `f400` and `f675`. All of the three variables exhibited very strong pairwise correlations among each other (Pearson's correlation of above 0.98). Variable `f282` was selected for the final model, as this variable had the strongest correlation with the observed losses among the three (approximately 0.332).

Six different types of generalized linear models were applied on the development data. Firstly, three models from Gamma family: one with identity link ($g(x) = x$), one with inverse link ($g(x) = \frac{1}{x}$) and one with log-link ($g(x) = \ln(x)$). Secondly, two models from inverse Gaussian family: one with identity link and one with log-link. Lastly, a model from lognormal family with identity link. The goodness of fit of the candidate models was assessed using Akaike information criterion (AIC) [36]. AIC for a particular model can be calculated as

$$AIC = -2 \cdot \ln(L) + 2 \cdot p,$$

where L represents the maximum value of the likelihood function and p represents the number of model parameters. As the models were estimated in R using the function `glm`, the value of AIC obtained for the lognormal models needs to be adjusted in order to compare it to other candidate models. This is due to the fact that the log-likelihood function in R is applied by default on the natural logarithm of loss in case of the lognormal model. The issue can be mitigated by calculating the sum of $\ln(loss)$ across the validation sample and adding it to the AIC value that is originally estimated by R. The comparison of AIC for all the candidate models is presented in the table below.

Family	link	AIC
Gamma	identity	39 450
	inverse	39 437
	log	39 315
inverse Gaussian	identity	38 690
	log	38 627
lognormal	identity	38 387

Table 3.2: Akaike information criterion for the candidate models

As can be seen from the table above, the lowest AIC was observed for the lognormal model, which is why this model was chosen as the final model. Using the lognormal model, the expected value of $\ln(loss)$ can be expressed as

$$\mathbb{E}[\ln(loss)] = \hat{\beta}_0 + \hat{\beta}_1 \cdot x_1 + \hat{\beta}_2 \cdot x_2 = 0.589789 + 1.915026 \cdot f404 + 0.039608 \cdot f282.$$

In order to obtain the estimates of loss, the predictions have to be transformed back to the original scale and thus,

$$\mathbb{E}(loss) = \exp(0.58979 + 1.91503 \cdot f404 + 0.03961 \cdot f282 + \frac{1}{2} \cdot \hat{\sigma}^2),$$

where the dispersion parameter $\hat{\sigma}^2$ was estimated to be approximately 0.80169. As a next step, the model was applied on the validation sample to test the discriminatory power. For testing purposes, both the observed losses and the estimated losses were discretized into 11 segments based on the intervals proposed by the regulatory reporting templates [21]. The first loss (estimation) segment consists of observations where the observed loss (estimated loss) is below 5%. The second loss (estimation) segment consists of observation for which the observed loss (estimated loss) is greater than or equal to 5%, but lower than 10%. The remaining 9 segments are constructed in a similar manner, where each remaining segment covers a 10% interval. Thus, the last loss (estimation) segment consists of observations where the observed loss (estimated loss) is greater than or equal to 90%. It was noted that one of the observations in the validation sample obtained a loss estimate over 100% due to parameter $f282$ having a significantly higher value for this observation compared to the rest of the sample. The second highest loss estimate was approximately 32.5%. If the model estimates are capped at 100% then the average observed loss in the validation sample is approximately 8.57% and the average loss estimate is approximately 8.55%. The resulting contingency table after discretization is presented in the table below.

estimate/loss	0%-5%	5%-10%	10%-20%	20%-30%	30%-40%	40%-50%	50%-60%	60%-70%	70%-80%	80%-90%	90%+
0%-5%	418	112	39	8	1	2	0	1	0	0	0
5%-10%	895	533	245	75	13	10	7	2	0	0	2
10%-20%	221	210	219	84	38	19	10	7	3	5	8
20%-30%	14	10	20	13	5	4	0	0	0	1	4
30%-40%	1	0	0	1	0	0	0	0	0	0	0
40%-50%	0	0	0	0	0	0	0	0	0	0	0
50%-60%	0	0	0	0	0	0	0	0	0	0	0
60%-70%	0	0	0	0	0	0	0	0	0	0	0
70%-80%	0	0	0	0	0	0	0	0	0	0	0
80%-90%	0	0	0	0	0	0	0	0	0	0	0
90%+	1	0	0	0	0	0	0	0	0	0	0

Table 3.3: Contingency table of discretized model estimates and losses, validation sample

The most significant concentrations can be observed in the second estimation segment (which holds approximately 54.65% of all observations) and in the first loss segment (which holds approximately 47.53% of all observations). The discriminatory power of the model was assessed across 8 different metrics: Kendall's τ, Goodman and Kruskal's γ, Somers' D (both ways), gAUC, τ^b , τ^c and Spearman's ρ. It must be noted that a commonly used function for estimating correlations in R, `cor`, calculates the value of τ^b when using option `method = c("kendall")`, instead of Kendall's τ. The correlation coefficients presented in the current thesis were estimated using functions `KendallTauA`, `GoodmanKruskalGamma`, `SomersDelta`, `KendallTauB`, `StuartTauC` and `SpearmanRho` from the R library `DescTools` [32]. The results of the analysis are presented in the table below.

τ	0.2067
γ	0.4920
D (loss|estimate)	0.3413
D (estimate|loss)	0.3063
gAUC	0.6707
τ^b	0.3233
τ^c	0.2479
ρ	0.3629

Table 3.4: Assessment of model's discriminatory power across various metrics

When assessing the results in the table above, it must be noted that result of gAUC is not directly comparable to other measures as the values of gAUC range from 0 to 1, but the values of all other metrics range from -1 to 1. The value of gAUC is derived directly from Somers' D of observed loss with respect to estimated loss. The fact that the absolute value of γ is more than two times higher compared to the absolute value of τ indicates that the population contains a significant proportion of tied observations (across both variables), which is also visible from Table 3.3. The fact that γ obtains a value close to 0.5 means that the correlation among pairwise comparable observations is quite strong. The value of τ^c is relatively low compared to the value of τ^b due to the fact that in the estimation of τ^c, the number of tied pairs across the rows and columns of the contingency table (which are relatively high) are not accounted for.

In order to illustrate the sensitivity of the measures of discriminatory power in relation to the increase of concentration in the first loss segment, the border of the first loss segment was expanded. In the first iteration, the first loss interval was changed from $[0\%, 5\%)$ to $[0\%, 6.25\%)$ and the second loss interval was changed from $[5\%, 10\%)$ to $[6.25\%, 10\%)$. This way, the first loss segment holds approximately 60.56% of all observations. The resulting contingency table together with results across the measures of discriminatory power are presented in the tables below.

estimate/loss	0%-6.25%	6.25%-10%	10%-20%	20%-30%	30%-40%	40%-50%	50%-60%	60%-70%	70%-80%	80%-90%	90%+
0%-5%	490	40	39	8	1	2	0	1	0	0	0
5%-10%	1164	264	245	75	13	10	7	2	0	0	2
10%-20%	300	131	219	84	38	19	10	7	3	5	8
20%-30%	19	5	20	13	5	4	0	0	0	1	4
30%-40%	1	0	0	1	0	0	0	0	0	0	0
40%-50%	0	0	0	0	0	0	0	0	0	0	0
50%-60%	0	0	0	0	0	0	0	0	0	0	0
60%-70%	0	0	0	0	0	0	0	0	0	0	0
70%-80%	0	0	0	0	0	0	0	0	0	0	0
80%-90%	0	0	0	0	0	0	0	0	0	0	0
90%+	1	0	0	0	0	0	0	0	0	0	0

Table 3.5: Contingency table after the first shift of the lowest loss segment

τ	0.1975
γ	0.5343
D (loss|estimate)	0.3262
D (estimate|loss)	0.3372
gAUC	0.6631
τ^b	0.3317
τ^c	0.2370
ρ	0.3696

Table 3.6: Assessment of discriminatory power after the first shift of the lowest loss segment

The increase in pairwise tied observation (across either variable) is illustrated by the decrease in τ and τ^c. The fact that the proportion of tied losses is higher in the shifted view compared to the proportion of tied model estimates is visible from the comparison of the two values of Somers' D: Somers' D of observed loss with respect to estimated loss is lower than Somers' D of estimated loss with respect to observed loss. The increase in γ and Somers' D of estimated loss with respect to observed loss should be interpreted with caution as the number of pairwise comparable observations in the shifted view is lower than in the original view, meaning that the result has less statistical evidence. The same general principle applies for τ^b, where the denominator in the calculation formula (see Equality 2.11) decreases together with the nominator. The fact that γ increases means that the proportion of disconcordant pairs in the population (among comparable pairs) has decreased in relation to the proportion of concordant pairs.

Next, the concentration of observation in the first loss segment was further increased by merging the first two loss segments. This resulted in a rectangular contingency table where the concentration of observations in the first loss segment (covering an interval of $[0\%, 10\%)$) is approximately 74%.

estimate/loss	0%-10%	10%-20%	20%-30%	30%-40%	40%-50%	50%-60%	60%-70%	70%-80%	80%-90%	90%+
0%-5%	530	39	8	1	2	0	1	0	0	0
5%-10%	1428	245	75	13	10	7	2	0	0	2
10%-20%	431	219	84	38	19	10	7	3	5	8
20%-30%	24	20	13	5	4	0	0	0	1	4
30%-40%	1	0	1	0	0	0	0	0	0	0
40%-50%	0	0	0	0	0	0	0	0	0	0
50%-60%	0	0	0	0	0	0	0	0	0	0
60%-70%	0	0	0	0	0	0	0	0	0	0
70%-80%	0	0	0	0	0	0	0	0	0	0
80%-90%	0	0	0	0	0	0	0	0	0	0
90%+	1	0	0	0	0	0	0	0	0	0

Table 3.7: Contingency table after the second shift of the lowest loss segment

τ	0.1582
γ	0.5751
D (loss|estimate)	0.2613
D (estimate|loss)	0.3746
gAUC	0.6307
τ^b	0.3129
τ^c	0.1898
ρ	0.3409

Table 3.8: Assessment of discriminatory power after the second shift of the lowest loss segment

The tables above illustrate that compared to the first shift of loss segment, Somers' D of observed loss with respect to estimated loss has decreased significantly, which is in line with the increased concentration of observations in the first loss segment. For the same reason the values of τ, τ^c and gAUC have also decreased compared to the previous iteration. However, both γ and Somers' D of estimated loss with respect to observed loss have increased. Lastly, the concentration of observation in the first loss segment was increased for the third time by merging the first two loss segments from the previous shift of loss buckets. The results can be seen from the tables below.

estimate/loss	0%-20%	20%-30%	30%-40%	40%-50%	50%-60%	60%-70%	70%-80%	80%-90%	90%+
0%-5%	569	8	1	2	0	1	0	0	0
5%-10%	1673	75	13	10	7	2	0	0	2
10%-20%	650	84	38	19	10	7	3	5	8
20%-30%	44	13	5	4	0	0	0	1	4
30%-40%	1	1	0	0	0	0	0	0	0
40%-50%	0	0	0	0	0	0	0	0	0
50%-60%	0	0	0	0	0	0	0	0	0
60%-70%	0	0	0	0	0	0	0	0	0
70%-80%	0	0	0	0	0	0	0	0	0
80%-90%	0	0	0	0	0	0	0	0	0
90%+	1	0	0	0	0	0	0	0	0

Table 3.9: Contingency table after the third shift of the lowest loss segment

τ	0.0801
γ	0.6406
D (loss|estimate)	0.1323
D (estimate|loss)	0.4336
gAUC	0.5662
τ^b	0.2395
τ^c	0.0961
ρ	0.2565

Table 3.10: Assessment of discriminatory power after the third shift of the lowest loss segment

The concentration of observations in the first loss segment after the last shift is approximately 90.1%. As expected, the values of τ, τ^c, Somers' D of observed loss with respect to estimated loss and gAUC have all significantly decreased compared to previous iterations. In addition, the correlations estimated by both τ^b and Spearman's ρ have also deteriorated. The table below presents an combined overview of all the results that were observed.

% obs. in first loss segment	τ	γ	Somers' D of loss with respect to estimate	Somers' D of estimate with respect to loss	gAUC	τ^b	τ^c	ρ
47.53%	0.207	0.492	0.341	0.306	0.671	0.323	0.248	0.363
60.56%	0.198	0.534	0.326	0.337	0.663	0.332	0.237	0.370
74.06%	0.158	0.575	0.261	0.375	0.631	0.313	0.190	0.341
90.10%	0.080	0.641	0.132	0.434	0.566	0.240	0.096	0.257

Table 3.11: Overview of correlations across the four scenarios

Table 3.11 illustrates that the shifts of the lowest loss segments decrease the values of τ, τ^c, Somers' D of observed loss with respect to estimated loss and gAUC after each step. This decrease can be explained by the decrease in the proportion of concordant and disconcordant pairs in the sample, while the total number of comparisons made remains the same across all iterations. This means that the more severe the shift is, the closer the value of the four measures will be to 0. The values of τ^b and Spearman's ρ are quite similar to the originally calculated values after the first two shifts, but significantly decrease after the last shift. For τ^b, the decrease in the proportion of concordant and disconcordant pairs is partly mitigated by the decrease of the denominator in the calculation formula (see Equality 2.11). As stated above, the increase in the values of γ and Somers' D of estimated loss with respect to observed loss should be interpreted with caution due to the fact that with each iteration, the number of comparisons considered in the calculations decreases. The increase in both measures illustrates the fact that the model is in general able to distinguish higher losses from lower losses better across the sample than the observations situated in the lowest loss segments from each other. The table below provides an overview of the total number of unique comparisons performed for the calculation of quantities that are related to Kendall's τ. The number of comparisons for τ^b is approximated by the square root of the product of the comparisons made during the calculations of Somers' D.

% obs. in first loss segment	Total number of unique comparisons considered in the calculation					
	τ	γ	Somers' D of loss with respect to estimate	Somers' D of estimate with respect to loss	τ^b	τ^c
47.53%	5 315 430	2 232 890	3 218 507	3 586 001	3 397 289	4 430 884
60.56%		1 965 250		3 114 251	3 165 950	
74.06%		1 462 541		2 245 251	2 688 188	
90.10%		664 895		982 206	1 777 987	

Table 3.12: The total number of unique comparisons considered in the calculations

As a last step, the values of τ and Somers' D of observed loss with respect to estimated loss were examined in relation to the corresponding values of τ^{max} and D^{max} across each iteration.

% obs. in first loss segment	τ	τ^{max}	*Somers' D of loss with respect to estimate*	D^{max}
47.53%	0.207	0.420	0.341	0.694
60.56%	0.198	0.370	0.326	0.611
74.06%	0.158	0.275	0.261	0.454
90.10%	0.080	0.125	0.132	0.207

Table 3.13: The comparison of τ and Somers' D to their estimated maximum values

From the table above, it can be seen that given the proportion of tied observations across the samples, it would not have been possible to observe correlation values close to 1 even if all the non-tied pairs were concordant. In case of the last iteration, the maximum value of Somers' D barely exceeds 0.2, which indicates that the maximum value of gAUC for that sample is approximately 0.6.

Conclusions

It was shown in the thesis that the measure used for reporting on the discriminatory power of IRB LGD models, which is defined by a linear transformation of Somers' D and referred to as generalized AUC (gAUC), will indicate relatively neutral correlation on the validation sample if the sample holds an excessive concentration of observations with tied losses. According to the regulatory reporting templates, the measure should be calculated considering discretized loss proportions and it obtains values between 0 and 1. The more severe the concentration of tied losses in the sample is, the closer the value of the measure will be to $\frac{1}{2}$. One example of a situation where this could become relevant is the case of low-loss portfolios, where the proportion of observed loss is close to zero for a large portion of the observations in the sample.

Although not explicitly required by the regulation, it makes sense from a practical perspective for financial institutions with IRB permission to incorporate the test into their annual validation methodology, as those results have to be reported each year together with the validation results. This means that an assessment scale has to be developed for the interpretation of the test results. However, the reporting instructions do not define any thresholds for assessing the results of the underlying gAUC measure and focus on evaluating whether the discriminatory power of the model at the reporting date has deteriorated compared to the initial validation of the model.

If the volume of pairwise-comparable observations in the sample is deemed sufficient and the model performs adequately in all other aspects, the relatively low value of gAUC should not dismiss the developed model in situations where excessive concentrations of observations with tied losses are present in the validation sample. For linear models, the square of the Pearson's correlation coefficient may be interpreted as the proportion of variance in one variable that is accounted for by the differences in the other variable. This relationship can be used as a basis to set the borders of acceptable Pearson's correlation for linear models. However, such interpretation is not applicable when dealing with monotonic correlation measures such as gAUC. In general, the interpretation of gAUC depends on the circumstances (sample and / or objective) and the final decision in terms of model acceptance could be supported with additional investigation. This includes the analysis across alternative measures of discriminatory power, extension of the observation-horizon, qualitative examination of the data sample, expert-based review of the current and previous versions of the model and benchmarking against competitors or challenger models.

Two supportive measures, the maximum value of Kendall's τ and the maximum value of Somers' D, given the concentration of tied pairs in the sample (τ^{max} and D^{max}) are introduced in the thesis, which can be used to assess the impact from having tied pairs of observations in the sample. It is also shown that if D^{max} is treated as an adjustment to Somers' D, then the adjusted correlation coefficient is equal to Goodman and Kruskal's γ. This indicates that γ could also be used as a supportive measure when assessing the results of gAUC during model validation.

References

[1] Bank for International Settlements. History of the Basel Committee. URL: https://www.bis.org/bcbs/history.htm (last visited: 31.07.2022).

[2] Basel Committee on Banking Supervision (1988). International convergence of capital measurement and capital standards. URL: https://www.bis.org/publ/bcbs04a.pdf (last visited: 31.07.2022).

[3] Basel Committee on Banking Supervision (1996). Amendment to the capital accord to incorporate market risks. URL: https://www.bis.org/publ/bcbs24.pdf (last visited: 31.07.2022).

[4] Basel Committee on Banking Supervision (2005). International Convergence of Capital Measurement and Capital Standards. URL: https://www.bis.org/publ/bcbs118.pdf (last visited: 31.07.2022).

[5] Basel Committee on Banking Supervision (2010). Basel III: International framework for liquidity risk measurement, standards and monitoring. URL: https://www.bis.org/publ/bcbs188.pdf (last visited: 31.07.2022).

[6] Basel Committee on Banking Supervision (2011). Basel III: A global regulatory framework for more resilient banks and banking systems. URL: https://www.bis.org/publ/bcbs189.pdf (last visited: 31.07.2022).

[7] Basel Committee on Banking Supervision (2013). Basel III: The Liquidity Coverage Ratio and liquidity risk monitoring tools. URL: https://www.bis.org/publ/bcbs238.pdf (last visited: 31.07.2022).

[8] Basel Committee on Banking Supervision (2014). Basel III: the net stable funding ratio. URL: https://www.bis.org/bcbs/publ/d295.pdf (last visited: 31.07.2022).

[9] Basel Committee on Banking Supervision (2017). Basel III: Finalising post-crisis reforms. URL: https://www.bis.org/bcbs/publ/d424.pdf (last visited: 31.07.2022).

[10] Brown, M. B. & Benedetti, J. K. (1977). Sampling Behavior of Test for Correlation in Two-Way Contingency Tables. *Journal of the American Statistical Association*, **volume 72**, pages 309-315.

[11] Committee on Banking Regulations and Supervisory Practices (1975). Report to the Governors on the supervision of banks' foreign establishments. https://www.bis.org/publ/bcbs00a.pdf (last visited: 31.07.2022).

[12] Committee on Banking Regulations and Supervisory Practices (1983). Principles for the supervision of banks' foreign establishments. https://www.bis.org/publ/bcbsc312.pdf (last visited: 31.07.2022).

[13] Cleff, T. (2014). *Exploratory Data Analysis in Business and Economics. An Introduction Using SPSS, Stata, and Excel*. Springer.

[14] European Banking Authority (2016). Draft Regulatory Technical Standards on the materiality threshold for credit obligations past due under Article 178 of Regulation (EU) No 575/2013. URL: https://www.eba.europa.eu/sites/default/documents/files/documents/10180/1597002/fe1db887-c6dc-4777-89c1-4f243584cafd/Final%20draft%20RTS%20on%20the%20materiality%20threshold%20for%20credit%20obligations%20%28EBA-RTS-2016-06%29.pdf?retry=1 (last visited: 31.07.2022).

[15] European Banking Authority (2016). Guidelines on the application of the definition of default under Article 178 of Regulation (EU) No 575/2013. URL: https://www.eba.europa.eu/sites/default/documents/files/documents/10180/1597103/004d3356-a9dc-49d1-aab1-3591f4d42cbb/Final%20Report%20on%20Guidelines%20on%20default%20definition%20%28EBA-GL-2016-07%29.pdf?retry=1 (last visited: 31.07.2022).

[16] European Banking Authority (2016). Final Draft Regulatory Technical Standards on the specification of the assessment methodology for competent authorities regarding compliance of an institution with the requirements to use IRB Approach in accordance with Articles 144(2), 173(3) and 180(3)(b) of Regulation (EU) No 575/2013. URL: https://www.eba.europa.eu/sites/default/documents/files/documents/10180/1525916/e8373cbc-cc4b-4dd9-83b5-93c9657a39f0/Final%20Draft%20RTS%20on%20Assessment%20Methodology%20for%20IRB.pdf?retry=1 (last visited: 31.07.2022).

[17] European Banking Authority (2017). Guidelines on PD estimation, LGD estimation and the treatment of defaulted exposures. URL: https://www.eba.europa.eu/sites/default/documents/files/documents/10180/2033363/6b062012-45d6-4655-af04-801d26493ed0/Guidelines%20on%20PD%20and%20LGD%20estimation%20%28EBA-GL-2017-16%29.pdf?retry=1 (last visited: 31.07.2022).

[18] European Banking Authority (2018). Final Draft Regulatory Technical Standards on the specification of the nature, severity and duration of an economic downturn in accordance with Articles 181(3)(a) and 182(4)(a) of Regulation (EU) No 575/2013. URL: https://www.eba.europa.eu/sites/default/documents/files/documents/10180/2459703/3136b895-0dfb-454f-8984-beddb888b8cc/EBA%20BS%202018%20xxx%20%28Final%20draft%20RTS%20on%20economic%20downturn%29_final%20%28002%29.pdf?retry=1 (last visited: 31.07.2022).

[19] European Banking Authority (2019). Final report on guidelines for the estimation of LGD appropriate for an economic downturn. URL: https://www.eba.europa.eu/sites/default/documents/files/documents/10180/2551996/f892da33-5cb2-44f8-ae5d-68251b9bab8f/Final%20Report%20on%20Guidelines%20on%20LGD%20estimates%20under%20downturn%20conditions.pdf?retry=1 (last visited: 31.07.2022).

[20] European Central Bank (2019). ECB guide to internal models. URL: https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.guidetointernalmodels_consolidated_201910~97fd49fb08.en.pdf (last visited: 31.07.2022).

[21] European Central Bank (2019). Instructions for reporting the validation results of internal models, IRB Pillar I models for credit risk. URL: https://www.bankingsupervision.europa.eu/banking/tasks/internal_models/shared/pdf/instructions_validation_reporting_credit_risk.en.pdf (last visited: 31.07.2022).

[22] European Parliament and the Council of the European Union (2013). Regulation (EU) No 575/2013 of the European Parliament and of the Council of 26 June 2013 on prudential requirements for credit institutions and investment firms and amending Regulation (EU) No 648/2012. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013R0575&from=en (last visited: 31.07.2022).

[23] European Parliament and the Council of the European Union (2019). Regulation (EU) 2019/876 of the European Parliament and of the Council of 20 May 2019 amending Regulation (EU) No 575/2013 as regards the leverage ratio, the net stable funding ratio, requirements for own funds and eligible liabilities, counterparty credit risk, market risk, exposures to central counterparties, exposures to collective investment undertakings, large exposures, reporting and disclosure requirements, and Regulation (EU) No 648/2012. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019R0876&from=EN (last visited: 31.07.2022).

[24] European Parliament and the Council of the European Union (2021). Proposal for a regulation of the European Parliament and of the Council amending Regulation (EU) No 575/2013 as regards requirements for credit risk, credit valuation adjustment risk, operational risk, market risk and the output floor. URL: https://eur-lex.europa.eu/resource.html?uri=cellar:14dcf18a-37cd-11ec-8daf-01aa75ed71a1.0001.02/DOC_1&format=PDF (last visited: 31.07.2022).

[25] Goodman, L. A & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association*, **volume 49**, pages 732-764.

[26] IFRS Foundation (2022). IAS 39 Financial Instruments: Recognition and Measurement. URL: https://www.ifrs.org/content/dam/ifrs/publications/pdf-standards/english/2022/issued/part-a/ias-39-financial-instruments-recognition-and-measurement.pdf?bypass=on (last visited: 04.08.2022).

[27] IFRS Foundation (2022). IFRS 9 Financial Instruments. URL: https://www.ifrs.org/content/dam/ifrs/publications/pdf-standards/english/2022/issued/part-a/ifrs-9-financial-instruments.pdf?bypass=on (last visited: 04.08.2022).

[28] Kaggle (2014). Loan default prediction - Imperial College London. URL: https://www.kaggle.com/competitions/loan-default-prediction/overview (last visited: 31.07.2022).

[29] Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, **volume 30**, pages 81-93.

[30] Kendall, M. G. (1970). *Rank Correlation Methods* (4th ed., in Russian). Griffin, London.

[31] Li, D., Bhariok, R. & Keenan, S. (2009). Validation techniques and performance metrics for loss given default models. *Journal of Risk Model Validation*, **volume 3**, pages 3-26.

[32] R Core Team. Documentation for package 'DescTools'. URL: `https://www.rdocumentation.org/packages/DescTools/versions/0.99.45` (last visited: 04.08.2022).

[33] R Core Team (2020). A software environment for statistical computing and graphics. URL: `https://www.r-project.org` (last visited: 04.08.2022).

[34] Rodgers, J. L. & Nicewander, W. A. (1988). Thirteen Ways to Look at the Correlation Coefficient. *The American Statistician*, **volume 42**, pages 59-66.

[35] Somers, R. H. (1962). A New Asymmetric Measure of Association for Ordinal Variables. *American Sociological Review*, **volume 27**, pages 799-811.

[36] Tutz, G. (2012). *Regression for Categorical Data*. Cambridge University Press.

Non-exclusive licence to reproduce the thesis and make the thesis public

I, Kristo Visk

1. grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright, my thesis
   Measures to assess the discriminatory power of loss given default models,
   supervised by Meelis Käärik.

2. I grant the University of Tartu a permit to make the thesis specified in point 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 4.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in points 1 and 2.

4. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Kristo Visk
18/08/2022