SURVEY OF CREDIT RISK MODELS IN RELATION TO CAPITAL ADEQUACY FRAMEWORK FOR FINANCIAL INSTITUTIONS

Poomjai Nacaskul *

* FSVP, Quantitative Models and Enterprise Analytics, Siam Commercial Bank PCL, Thailand

Abstract

This article (i) iterates what is meant by credit risks and the mathematical-statistical modelling thereof, (ii) elaborates the conceptual and technical links between credit risk modelling and capital adequacy framework for financial institutions, particularly as per the New Capital Accord (Basel II)'s Internal Ratings-Based (IRB) approach, (iii) proffer a simple and intuitive taxonomy on contemporary credit risk modelling methodologies, and (iv) discuses in some details a number of key models pertinent, in various stages of development, to various application areas in the banking and financial sector.

Keywords: Credit Risk, Default Probability, Capital Adequacy Framework, Credit Derivatives, Rating Transition, Default (Asset Value) Process/Intensity Models, Linear/Nonlinear Discriminant Analysis, Loss Distribution, Quantitative Modelling.

JEL: M41, M48, F55

1. INTRODUCTION

Whereas a dictionary definition of "credit risk" merely alludes to the possibility of the borrower not returning the amount of money borrowed from the lender when due, as per agreed upon at the time of lending, a regulatory perspective necessarily generalises the notion of "credit" to a large class of business activities that financial institutions operate. As a corollary, the notion of "credit risk" also generalises to encompass all kinds of risk phenomena associated with said business.

This article properly begins by defining credit risks from three distinct lines of banking businesses that cradle around three distinct lines of financial products: loans, bonds, and derivatives.

1.1 What is Credit Risk?

Credit Risks, along with Market/Price Risks and Operational Risks, comprise the main types of Financial Risks, by which is meant risks that are (i) measured in monetary units, (ii) arisen from/rooted in financial-market variables/financial-institution factors, (iii) managed by means of financial techniques/tools, and/or (iv) seen as being intrinsic/integral to the system of financial markets/institutions.

In and of itself, credit risk refers to the possibility and probability of financially relevant losses arising from credit events, which in turns are defined in general terms thus:

For bank loans – credit risks cover all of the followings:

- Default risk pertains to the risk of significant losses due to failed principal and/or interest payments, whence including in this definition counterparty risk, i.e. the possibility of one's failing to honour transactional promises, thereby leading to significant opportunity losses arising from attempts to seek substitute transactions, which in some cases may not be possible. In other words, by letting both defaults on payments and such counterparty failures be referred to collectively as default events, default risk is seen as the specialisation of credit risks by virtue of the fact that whole of default events forms a strict subset within the universe of possible credit events. For modelling purposes, default events are most naturally captured by a random variable (r.v.) with a Bernoulli distribution, with Probability of Default (PD) as the one and only (distributional) parameter. In a sense, the determination of a default event occurring is akin to tossing a highly-skewed coin, with the probability of landing 'Head' (for default) much smaller than 1 in 2, with most cases landing 'Tail' (for non-default):

$$D \sim Bernoulli\left(p_{Default}\right) \Rightarrow \begin{cases} D \in \{0 \Leftrightarrow 'non\text{-}default', 1 \Leftrightarrow 'default'\} \\ \Pr('default') \equiv p(1) = p_{Default} \\ \Pr('non\text{-}default') \equiv p(0) = 1 - p_{Default} \\ \mathrm{E}[D] = 0 \cdot p(0) + 1 \cdot p(1) = p_{Default} \\ Var(D) \equiv \mathrm{E}\left[\left(D - \mathrm{E}[D]\right)^2\right] = p_{Default}\left(1 - p_{Default}\right) \end{cases} \quad (1)$$

- Collateral risk involves uncertainty in the realisable value of collateral assets in the event of default actually occurring. In a quantitative model, it is captured as Loss Given Default (LGD) in various guises, namely a simple numerical constant, a means parameter governing the distribution of some random variable, a function of time or other variables, or itself a random variable, etc., depending on the level of details and modelling methodology.

- Drawdown risk arises in connections with contingent claim products and open credit lines. It is captured with the quantity Exposure At Default (EAD) that, as with LGD, may be specified simply as a numerical constant or otherwise.

For bonds/fixed-income securities/debt instruments, whereupon there exist independent credit rating agencies who rate the creditworthiness of various debt issues and their issuers on a finite alpha-numeric ordinal scale, credit risks refer to downgrade risk, as every time a credit rating agency announces a 'credit downgrade', replacing one credit-rating grade with another a 'notch' or more lower down, or even a simple 'credit watch/alert', the implicated securities immediately suffer loss in value, as they are marked-to-market with new discount rates, ones reflecting yet higher credit spreads, i.e. discount rate differentials above (credit) risk-free bonds. Note that within this context, default event corresponds to the bottommost credit grade. Also, in this context a credit upgrade resulting in a marked-to-market gain in value may be interpreted in terms of upside risks.

For credit derivatives including credit default swaps (CDS), credit linked notes (CLN), collateralised debt obligations (CDO), and structured products of various types, credit risks refer to financial risks resulting from credit derivative deals and transactions, whereupon the exact nature of credit events depend on how they are stipulated in the derivatives contracts, so as to include, among the possibilities, single-name/multi-name bond default/downgrade, accounting/technical insolvency, business/corporate bankruptcy, debt/payment restructuring, and so on.

1.2 What is a Credit Risk Model?

Quantitative model refers to a "toy model" endowed with critical a number of mathematical properties obtained from real-life entities or systems, whereby an analyst may experiment with different details and configurations of various model components in order to engender in-depth understanding of the mechanisms, processes, as well as behaviours of the real item or system under realistic situations, especially where there are technical, time, cost, and/or ethical constraints prohibiting experimentation with real-life objects.

Financial risk management process comprises 4 main steps, namely to: 1. IDENTIFY/define, 2. MEASURE/assess, 3. MITIGATE/manage, and 4. REVIEW/report the pertinent financial risks. Contemporary risk management concepts abide strictly by the principle that any risk that can be mitigated must be quantitatively measured first.

Credit-risk models in this context therefore refer simply to quantitative models employed toward the task of measuring the extents of credit risks concerned. Nowadays, credit risk modelling is almost always based on probability theory and principally utilises statistical tools for performing analyses, so that probability distribution assumptions and parametric estimation techniques are often required.

In terms of multiplicity of the financial obligors involved, both varieties or credit risk models, namely single-obligor credit-risk models (for individual loans/bond issues and single-name credit derivatives) as well as multiple-obligor credit-risk models (for bank loan/bond investment portfolios and multi-name credit derivatives), exist and play complementary roles.

On the other hand, contemplating along the line of how analytical results are put to use, credit risk models may be categorized into 6 application areas/groups, namely:

Credit Determinant Models are single-obligor credit risk models used mostly in the analysis of individual bank loans, especially in the retail lending line of business:

- Beginning with the development of Credit Scoring Model (Altman, 1968), which, having empirically determined a set of relevant obligor attributes and corresponding factor-loading coefficients, then produces real-valued a credit score to be used in discriminating (potential) loan clients on the basis of creditworthiness:

$$a \mapsto s\,, \quad a \in AttributeSpace \subseteq \Re^n\,, \quad s \in RawScoreRange \subseteq \Re \quad (2)$$

- From this develops Loan Decision Model which compares a raw credit score against some specified threshold values and hence transforms it into a decision variable, such as whether to approve a particular loan as it stands, request more information and/or modify terms of lending, and so on:

$$s \mapsto l\,, \quad l = \begin{cases} "AutoApproved" & s \geq \theta_{high} \\ "MayBeApprovedIf..." & \theta_{low} \leq s < \theta_{high} \\ "AutoDeclined" & s < \theta_{low} \end{cases} \tag{3}$$

- And with sufficient accumulated data on historical defaults/non-defaults, it is possible to pursue a Statistical PD Estimation Model, for instance by converting the signal output from a credit scoring model into a numerical PD estimate, whose value by definition ranges from 0 (certain not to default) to being 100% certain of a default event:

$$a \mapsto s \mapsto \hat{p}\,, \quad \hat{p} \in [0,1] \tag{1}$$

In any event, whether the aim is credit scoring, loan decision, or statistical PD estimation, the ex post model performance depends on the inherent discriminatory power of the statistical inference engine that lies of the heart of any such methodology once applied to real-application data.

Rating Transition Models are single-obligor credit risk models used mostly in the analysis of rated corporate bonds, especially those that trade in secondary markets with sufficient liquidity, often relying on Discrete-/Continuous-Time Markov Chain (MC) as the key theoretical constructs. Such models create a set of transition probabilities from one credit rating grade (on an ordinal, alpha-numeric scale) to another, the whole of which is summarised into a transition probability matrix. As "D for default" constitutes the lowest possible grade, PD estimates are included in the corresponding bottom row of the matrix. Such a matrix also forms the basis from which Credit Spreads/Term Structure Models are derived.

Default Process Models are single-obligor credit risk models used mostly in the analysis of individual bank loans, especially large corporate loans. Here default event is modelled with some stochastic process, from which PD estimates may then be derived. Default process models come in two fairly distinct types, namely:

- Structural Default (Asset-value) Models set out to model asset value process $\{A_t, t \geq 0\}$ rather explicitly, whence over any horizon a default arises whenever the obligor's balance sheet got so structurally weak so as to be unable to continue operating as a going concern, the simplest case being one of insolvency, i.e. $\exists\, s \in (0,t] \ni A_s \leq 0$, whence PD is derived from the 'risk-neutral-world' probability associated with such an event.

- Default Intensity (Reduced-form) Models set out to model (stochastic) default arrival process $\{D_t, t \geq 0\}$ directly, with nothing to say about the obligor's balance sheet structure or any other types of internal information. Only a quantity called default intensity is relevant for the specification of the default arrival process, whereupon this default intensity itself may be specified as a constant $\lambda > 0$, a function of time, $\lambda(t), t \geq 0$, or even a stochastic process $\{\lambda_t, t \geq 0\}$, depending on the modelling concept and capability.

Credit Portfolio Models are multiple-obligor credit risk models used in the analysis of bank loan/bond investment portfolios, enlisting such theoretical developments and modelling tools as Bernoulli mixture models and Copula dependency models. Such techniques are also relevant in the analysis of double default risk (when an obligor with credit guarantor defaults, it's also possible for the credit guarantor, now taking the original obligor's place, to also default), PD-LGD correlation (when LGD is a random quantity that correlates statistically with the default event), and credit-market risk correlation (when credit events and market events are not completely independent). Often, credit portfolio models involve two-stage randomness, i.e. one randomness accounts for the number of credit events, each of which gives rise to another randomness associated with individual losses, where summing up individual losses generally calls for a mathematical concept/technique of convolution.

Moreover, there are Credit Derivatives Models (Chaplin, 2005; Meissner, 2005) used in the financial engineering process of designing, conducting research and development, productionising and pricing, i.e. calculating the financial cost of implementing a requisite hedging programme for all outstanding credit derivatives transactions/deals.

Finally there are Model Validation Tools, which are not models or modelled results in the above sense, but refer to analytical tools used in reviewing/testing/validating a model in terms of being theoretically sound as well as empirically supported (forecast errors, goodness of fit, etc.). As such, model validation tools find their uses more in middle-office/business-control applications than in front-end/business-line applications.

1.3 Credit Risk Models in relation to Capital Adequacy Framework

Risk Capitals, Economic Capitals & Capital Adequacy Framework – Risk Capitals within the context of credit risks in banking refers to the level of equity capital deemed sufficient to absorb financial losses from all realised credit events without affecting the ability to honour financial obligations, stipulated at some statistical confidence level deemed sufficient in the eyes of supervisory authority in her capacity as fiduciary over public deposits and credit rating agencies in their capacity as monitoring agents on behalf of principals, i.e. those who lend to financial institutions and/or hold papers issued by them.

Economic Capitals (EC) refers to risk capitals estimated at a given statistical level of confidence, which depends on equity shareholders' requirements (as indicated by the cost of equity capitals) and bondholders' requirements (as indicated by the cost of debt capitals). For public companies, determining EC involves an additional consideration, namely agency credit rating, i.e. should a public company wish to continue its 'AA' rating, and so on.

Capital Adequacy Framework for safeguarding bank assets' exposure to credit risks refers to the standard and measure in specifying the minimum level of bank capital, i.e. the Regulatory Minimum Capital (RC), that serves as a protection of general deposit accounts from financial losses arisen from credit events, stipulated at some statistical confidence level established by the regulatory authority and enforced by the supervisory agency, taking into account the natural trade-off between the will to protect the public and the desire to ease compliance costs so that banks remain operationally profitable. At any rate, RC specification cites conservatism as the guiding principle; whereas EC estimation cites accuracy as the guiding principle.

Obtaining EC/RC from the Credit Portfolio's Loss Distribution – despite the differences, EC and RC both correspond to quantities that may be determined given the statistical Loss Distribution of a Credit/Credit-Risk Sensitive Portfolio $L_{CreditPortfolio} \sim PortfolioCreditLossDistr$:

$$F(l) \equiv \Pr\left(L_{CreditPortfolio} \le l\right) = \int_{-\infty}^{l} f(s)\,ds = \int_{0}^{l} f(s)\,ds \quad \because f(l)\begin{cases} \ge 0 & l > 0 \\ = 0 & l < 0 \end{cases} \tag{5}$$

Begin with the definition of Expected Loss (EL), set to coincide with the means of said loss distribution:

$$ExpectedLoss: EL \equiv \mathrm{E}\left[L_{CreditPortfolio}\right] = \int_{-\infty}^{\infty} s \cdot f(s)\,ds \tag{6}$$

Being non-decreasing, the generalised inverse—which is the same as the usual function inverse when the distribution function is continuous and strictly increasing over the distributional support - of such a loss distribution exists and is called the quantile function:

$$QuantileFunction: Q(p) \equiv F^{\leftarrow}(p)\,, \quad p \in [0,1] \tag{7}$$

From which:

$$CreditVaR_{@100(1-\alpha)\%confidence} \equiv \begin{cases} F^{\leftarrow}(1-\alpha)\,, & or \\ F^{\leftarrow}(1-\alpha) - \mathrm{E}\left[L_{CreditPortfolio}\right] \end{cases} \tag{8}$$

Now EC - whence ideally RC as well - can be estimated in terms of Credit Value-at-Risk (Credit-VaR) calculation, taken at the $100(1-\alpha)\%$ statistical confidence level:

$$RC_{CreditRisk} \approx EC_{CreditRisk} \equiv F^{\leftarrow}(1-\alpha) - EL \tag{9}$$

Moreover, Unexpected Loss (UL) may be defined as the standard deviation (SD) of the loss distribution:

$$\begin{aligned} LossVariance: Var(L) &\equiv \mathrm{E}\left[\left(L - \mathrm{E}[L]\right)^2\right] = \int_{-\infty}^{\infty} (s - EL)^2 \cdot f(s)\,ds \\ UnexpectedLoss: UL &\equiv StdDev(L) = \sqrt{Var(L)} \end{aligned} \tag{10}$$

So that Credit-VaR can also be stated as an $m > 0$ multiple of UL:

$$\exists\, m > 0 \ni F^{-1}(1-\alpha) - EL = m \cdot UL \tag{11}$$

Credit Risk Models Capable of Generating Credit Portfolio's Loss Distribution – Clearly, of all the 6 application areas or model groups, only credit portfolio models (group 4) are capable of generating loss distribution as well as EC/RC information at the credit portfolio level. And in practice, any modelling framework capable of estimating EC could form a basis for specifying RC, any qualitative differences being minimal. As a matter of course, EC and RC do differ in quantitative terms, with higher statistical confidence level for the latter but more frequent updating required for the former. And so technological developments with regards to EC estimation techniques can be seen to progress along side attempts to align RC specification concepts with EC results.

Note how the first 3 model groups (credit determinant, rating transition, and default process models) at best only produce parametric information of loss distribution at the individual Credit Item/Credit-Risk Sensitive Asset level $L_{IndividualCredit} \sim IndividualCreditLossDistr$, thereby denying the degree of representation and realism required to serve as a foundation for the capital adequacy framework for financial institutions. In fact, it's generally the case that:

$$L_{CreditPortfolio} < \sum L_{IndividualCredit} \quad (12)$$

But then again, if it is preferable to accept a risk model with the property that should it err, it should err on the conservative side, then it would practical to accept, as a matter of definition:

$$\hat{L}_{CreditPortfolio} \equiv \sum L_{IndividualCredit} \quad (13)$$

1.4 Credit Risk Models in relation to Basel II

At present, contemporary with the Basel Committee on Banking Supervision (BCBS)'s release of New Capital Accord (Basel II), live implementations of credit portfolio models capable of generating credible EC/RC figures are few and far between, with the majority of credit risk modelling tools still devoted to single-obligor PD estimation, never mind a complete loss distribution (for PD is but a component of EL, and unless EAD and LGD are simple numerical constants, PD alone says insufficiently little about the value of UL nor the type of statistical distribution that should be used).

Taking all the above into considerations, it is no wonder that the heart of Basel II, in particular the application of Internal Ratings-Based (IRB) Approach in determining capital adequacy with respect to credit risk exposures, largely concerns taking results from single-obligor credit risk models already in use with financial institutions and translating them into credit risk contribution to EC estimation, that is, by way of a certain device, namely the (Basel II IRB's) Risk-Weight Functions (RWF)

Summarily, RWF was designed specifically to overcome 2 modelling constraints:

1. To do away with the need to develop a fully-fledged credit risk portfolio models, as only single-obligor credit models are widely available.
2. To go ahead and forge EC/RC/UL/Credit-VaR calculation even though no 2nd-/higher-order moments information is available, as PD estimation models can only estimate PD (which when multiplied expected EAD/LGD can only produce 1st-moment statistics)

That RWF is able to overcome these 2 modelling constraints is down to a theoretical construct named Asymptotic Single Risk Factor (ASRF), whereby:

1. The 'asymptotic' part of the assumption enables multiple-obligor credit risk to be calculated simply by adding together "correlation-adjusted" single-obligor credit risk *measures*, in other words without having to recomputed credit portfolio risk every time a credit item is introduced to the portfolio, due to credit correlation effects, which would otherwise render such an *incremental* calculation highly erroneous. This is exactly what is referred to as the special "portfolio invariant" property of the risk measure.
2. The 'single risk factor' part of the assumption enables the synthesis of 2nd-moment statistics (UL) from 1st-moment statistics (EL), that is, by transforming "unconditional EL" into "conditional EL", wherein instead an additional parametric assumption known as *factor correlation* had been introduced vis-à-vis single-obligor credit quality and systematic factor signifying economic duress.

In any event, even though the credit risk aspect of the capital adequacy framework really calls for credit portfolio models (4th application area credit risk models) as the basis for calculating RC, etc., due to practical and technological limitations, i.e. the fact that credit portfolio modelling have not "matured" in the same way that loan discriminant, rating transition, and default process models have, Basel II stipulates thus:

$$\begin{aligned} EC_{CreditPortfolio} &\equiv \sum EC_{IndividualCredit}, \\ EC_{IndividualCredit} &= RWF\left(PD_{IndividualCredit}\right) \end{aligned} \quad (14)$$

2. SURVEY

This section shall survey core developments in credit risk modelling with respect to the 4 main application areas, namely loan discriminant, rating transition, default process, and credit portfolio models, leaving out of the discussion credit derivatives models and various model validation tools, progressing along the following credit risk model taxonomy:

1. Credit Determinant Models
    a. Discriminant Analysis–based
        i. Linear Discriminant Analysis – linearly separable feature space
        ii. Support Vector Machine – nonlinearly separable feature space
    b. Regression Analysis–based
        i. Binary (Logit/Probit) Regression – linear, parametric estimation/classification
        ii. Artificial Neural Networks – nonlinear, semi-parametric estimation/classification
2. Rating Transition Models
    a. Discrete-Time Finite-State Transition
        i. Stationary Markov Chain (MC)
        ii. Nonstationary/Time-Heterogeneous MC
        iii. Non-Markov Process (w/ Persistence of Memory)
    b. Continuous-Time Finite-State Transition
        i. Continuous-Time Markov Process
        ii. Stochastic Transition Intensity Model
3. Default Process Models
    a. Structural Default (Asset-value) Models

i. Merton's Asset-value Model
ii. Black & Cox's First-Passage Model
iii. PD Calibration vs. Historical Default Data

b. Default Intensity (Reduced-form) Models
i. Forward Default Intensity/Hazard Rate Model
ii. Doubly Stochastic/Stochastic Default Intensity Model

4. Credit Portfolio Models
a. Default/Rating Transition Correlation Approaches
i. Bernoulli Mixture Approach
ii. Multivariate Normal Approach
iii. Distributional Copula Approach

b. Stochastic Arrival/Loss Convolution Approaches
i. Poisson/Renewal Arrival Process
ii. Mixed Poisson/Negative Binomial Counting Process
iii. Extreme-value Losses/Sub-exponential/Heavy-tailed Distributions

2.1 Credit Determinant Models

2.1.1 Discriminant Analysis - based – Linear Discriminant Analysis (LDA) (Fisher, 1936; Duda, Hart, and Stork, 2000) posits that there are two distinct subpopulations of attribute vectors, all distributed according to a Multivariate Normal Distribution with the same variance-covariance matrix $\Sigma \in \Re^{n \times n}$, but with either of two different mean vectors $\mu_0 \neq \mu_1$, $\mu_0, \mu_1 \in \Re^n$, one for each subpopulation. For credit risk application, the two subpopulations comprise those that 'will' default, identified by the indicator (variable) $y = 1$, and those that 'will not' default, identified by $y = 0$, thus:

$$A|_{y=0} \sim \mathrm{N}(\mu_0, \Sigma_0) \,\&\, A|_{y=1} \sim \mathrm{N}(\mu_1, \Sigma_1) \ni \Sigma_0 = \Sigma_1 = \Sigma \tag{2}$$

From this, it emerges that, for each obligor's attribute vector $A \in \{a\} \subseteq \Re^n$, it is possible to estimate PD from the computation of matrix inverse $\Sigma^{-1} \ni \Sigma^{-1}\Sigma = \mathrm{I}$ as such:

$$PD(a) \equiv \Pr(y = 1 \mid a) \quad \propto \quad a^T \beta_{LDA}, \quad \beta_{LDA} \equiv \Sigma^{-1}(\mu_1 - \mu_0) \tag{3}$$

With sufficient amount of empirical default and non-default data, $Data_{Default} \equiv \left\{ \left(a|_{y=1} \right)_j \right\}_{j=1}^{n}$ and $Data_{NonDefault} \equiv \left\{ \left(a|_{y=0} \right)_i \right\}_{i=1}^{m}$, it would be possible to use the former to estimate $\hat{\mu}_1 \ni \mathrm{E}[\hat{\mu}_1] = \mu_1$, the latter to estimate $\hat{\mu}_0 \ni \mathrm{E}[\hat{\mu}_0] = \mu_0$, and both datasets to estimate $\hat{\Sigma} \ni \mathrm{E}[\hat{\Sigma}] = \Sigma$, from which it is then possible to compute $\hat{\beta}_{LDA} = \hat{\Sigma}^{-1}(\hat{\mu}_1 - \hat{\mu}_0)$ as a basis for classifying new obligors.

2.1.2 Regression Analysis-based – Binary (Logit/Probit) Regression similarly relies on the existence and quality of past data, but instead determine regression coefficients $\hat{\beta}_{Logit}$ or $\hat{\beta}_{Probit}$ by performing the following sum of squared errors (SSE) minimisation:

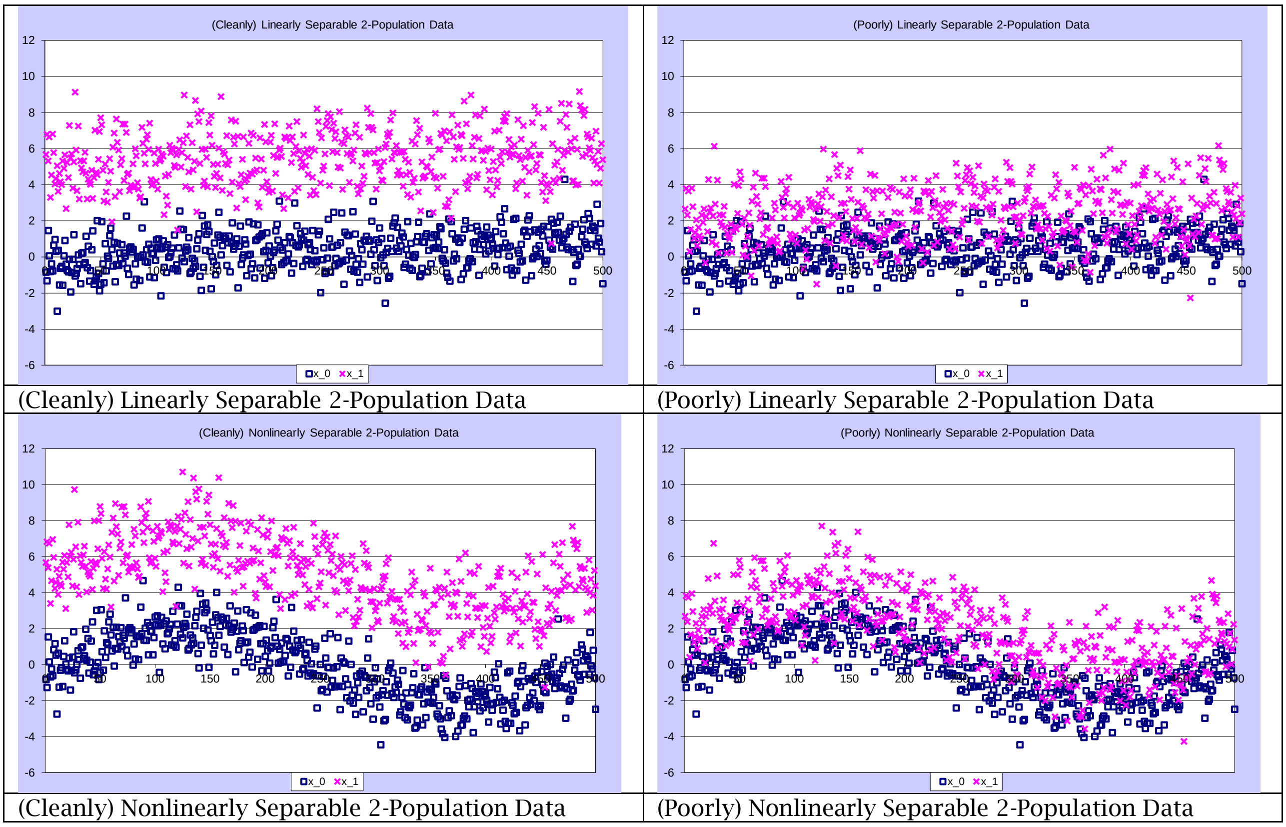
$$\min_{\beta \in \Re^n} \left\{ SSE \equiv \sum_{k=1}^{m+n} \left(y_k - \wp_{Logit/Probit}\left(\hat{\beta}^T_{Logit/Probit} a_k \right) \right) \right\}$$
$$, where \quad \wp_{Logit}(s) = \frac{1}{1 + e^{-s}} \in [0,1] \tag{4}$$
$$, whereas \quad \wp_{Probit}(s) = \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{s} e^{-x^2/2} dx \in [0,1]$$

Note how LDA and binary regression both rely on finding the coefficient vector with which to perform a scalar product $s = \beta^T a$, $\beta \in \{\beta_{LDA}, \beta_{Logit}, \beta_{Probit}\}$ with the attribute vector. All subsequent results are either monotonically increasing/non-decreasing or monotonically decreasing/non-increasing in $s \in \Re$, but never both.

This raises two technical issues: scaling and linearity. Of these, scaling is the simpler issue, and can be dealt with by a judicious choice of pre-processing procedures, often requiring straight-forward transformation of variable, such as from the binary 'sex' indicator into a Boolean variable $sex \in \{0,1\}$, or from 'age' in years to an ordinal-scale $age\_bucket \in \{1,2,...,7\}$, and so on.

The real limitation with LDA and binary regression is the fact that $\beta^T a$ is an essentially linear expression in the vector-valued variable, so that they can only work if it is possible/acceptable to draw a linear hyper-surface (generalization of a line in multi dimensions) delineating populations in their attribute (vector) space. In other words, they can only solve Linearly Separable Classification Problems. Dealing with Nonlinearly Separable Classification Problems (**Ошибка! Источник ссылки не найден.**) calls for a higher level of modelling sophistication, respectively, from LDA to Support Vector Machine (SVM) (Boser, Guyon, and Vapnik, 1992), and from binary regression to Artificial Neural Networks (ANN) (Pao, 1989).

Figure 1. Linearly vs. Nonlinearly Separable 2-Population Data



(Cleanly) Linearly Separable 2-Population Data | (Poorly) Linearly Separable 2-Population Data

(Cleanly) Nonlinearly Separable 2-Population Data | (Poorly) Nonlinearly Separable 2-Population Data

2.2 Rating Transition Models

2.2.1 Discrete-Time Finite-State Transition – For debt issues and issuers with agency credit ratings, the analysis of downgrade risk (which subsumes default risk as a special case) begins with a 1-year transition probability matrix, i.e. with $g = 8$ grades (7 non-defaults + 1 default) $\{AAA, AA, A, BBB, BB, B, CCC, Default\}$, use $\Omega \in \Re^{8\times 8}$:

$$\Omega = \begin{pmatrix} 90 & 1 & 0.1 & 0 & 0 & 0 & 0 & 0 \\ 8 & 90 & 6 & 0.1 & 0.1 & 0.1 & 0 & 0 \\ 1 & 7 & 85 & 6 & 0.2 & 0.2 & 0.2 & 0 \\ 0.7 & 1 & 7 & 85 & 9 & 0.7 & 0.8 & 0 \\ 0.2 & 0.7 & .9 & 7 & 80 & 9 & 5 & 0 \\ 0.1 & 0.2 & 0.7 & 1 & 9 & 75 & 9 & 0 \\ 0 & 0.1 & 0.2 & 0.7 & 1 & 10 & 65 & 0 \\ 0 & 0 & 0.1 & 0.2 & 0.7 & 5 & 20 & 100 \end{pmatrix},$$

$$\langle \Omega \rangle_{ij} = \Pr(Transition to Grade\, i \;\; from Grade\, j),$$

$$i, j \in \left\{ \underbrace{AAA}_{1}, \overbrace{AA}^{2}, \underbrace{A}_{3}, \overbrace{BBB}^{4}, \underbrace{BB}_{5}, \overbrace{B}^{6}, \underbrace{CCC}_{7}, \overbrace{Default}^{8} \right\} \quad (5)$$

To work with such a transition probability matrix, devise a ***rating state vector***:

$$x \in \left\{ \underbrace{\begin{pmatrix}1\\0\\0\\0\\0\\0\\0\\0\end{pmatrix}}_{AAA}, \overbrace{\begin{pmatrix}0\\1\\0\\0\\0\\0\\0\\0\end{pmatrix}}^{AA}, \underbrace{\begin{pmatrix}0\\0\\1\\0\\0\\0\\0\\0\end{pmatrix}}_{A}, \overbrace{\begin{pmatrix}0\\0\\0\\1\\0\\0\\0\\0\end{pmatrix}}^{BBB}, \underbrace{\begin{pmatrix}0\\0\\0\\0\\1\\0\\0\\0\end{pmatrix}}_{BB}, \overbrace{\begin{pmatrix}0\\0\\0\\0\\0\\1\\0\\0\end{pmatrix}}^{B}, \underbrace{\begin{pmatrix}0\\0\\0\\0\\0\\0\\1\\0\end{pmatrix}}_{CCC}, \overbrace{\begin{pmatrix}0\\0\\0\\0\\0\\0\\0\\1\end{pmatrix}}^{Default}, \right\} \tag{6}$$

It's then possible to extract 1-year rating probability vector from any current rating grade, for example 'BBB':

$$P_{BBB}^{(1)} = \Omega x_{BBB}$$
$$= \begin{pmatrix} 90 & 1 & 0.1 & \underline{0} & 0 & 0 & 0 & 0 \\ 8 & 90 & 6 & \underline{0.1} & 0.1 & 0.1 & 0 & 0 \\ 1 & 7 & 85 & \underline{6} & 0.2 & 0.2 & 0.2 & 0 \\ 0.7 & 1 & 7 & \underline{85} & 9 & 0.7 & 0.8 & 0 \\ 0.2 & 0.7 & .9 & \underline{7} & 80 & 9 & 5 & 0 \\ 0.1 & 0.2 & 0.7 & \underline{1} & 9 & 75 & 9 & 0 \\ 0 & 0.1 & 0.2 & \underline{0.7} & 1 & 10 & 65 & 0 \\ 0 & 0 & 0.1 & \underline{0.2} & 0.7 & 5 & 20 & 100 \end{pmatrix} \begin{pmatrix}0\\0\\0\\\underline{1}\\0\\0\\0\\0\end{pmatrix} = \begin{pmatrix}0\\0.1\\6\\85\\7\\1\\0.7\\0.2\end{pmatrix} \tag{7}$$

Moreover, in practice suppose there were a set of 1-year forward term structures of interests, one for each credit-rating grade (all except 'Default'), computing a credit risk-adjusted bond pricing could be done in a straightforward manner. First discount all future cashflows according to the 1-year forward rates for each credit-rating grade. Then multiply the forward prices corresponding to each grade with the probability of transiting to that grade and add up the weighted sum (for 'Default', substitute recovery value instead), for instance:

$$Fwd.PriceOf\_BBB\text{–}ratedBond =$$
$$\begin{pmatrix}0 & 0.1 & 6 & 85 & 7 & 1 & 0.7 & 0.2\end{pmatrix} \begin{pmatrix} Fwd.PriceIf \mapsto AAA \\ Fwd.PriceIf \mapsto AA \\ Fwd.PriceIf \mapsto A \\ Fwd.PriceIf \mapsto BBB \\ Fwd.PriceIf \mapsto BB \\ Fwd.PriceIf \mapsto B \\ Fwd.PriceIf \mapsto CCC \\ Recov.ValueIfDefault \end{pmatrix} \tag{8}$$

Moreover, should such a transition probability matrix remain constant from one year to the next, a stochastic-process property known as stationarity (particularly relevant for the through-the-cycle credit-rating regime), rating probability vectors can be calculated beyond just 1 year thus:

$$P_g^{(n)} = \Omega P_g^{(n-1)} = \Omega^2 P_g^{(n-2)} = \cdots = \Omega^n x_g \ , \quad g \in \{AAA, AA, A, BBB, BB, B, CCC, Default\} \tag{9}$$

All these theoretical/computational niceties subsume under a Stationary (Finite-State) Markov Chain (MC) model. On the other hand, should it be preferable to reflect short-term dynamics (particularly relevant for the point-in-time credit-rating regime), then it becomes necessary to give up the stationarity property, thereby allowing the transition probability matrix, not merely to update, but to change meaningfully from one year to the next. Moreover, the Markov (Memorylessness) property may not hold in reality, such as when empirical studies suggest that debt issues/issuers sharing the same current grade may transit according to different transition probability vectors, depending on last year's rating grades, therefore distinguishing those who came from a higher grade (credit deterioration), those who remain in the same grade, and those who came from a lower grade (credit improvement). This requires modelling with Non-Markov/Process with Memory.

2.2.2 Continuous-Time Finite-State Transition – Note how MC is akin to a situation whereby credit ratings are adjusted all at once, once a year. But in reality, credit-rating agencies change rating grades whenever the circumstances warrant, so that a more realistic downgrade risk model should be based on Continuous-Time Markov Process, where instead of citing the transition probability matrix $\Omega \in \Re^{g \times g}$ as the starting point of analysis, the modelling process would begin with the specification of the parametric transition intensity matrix $\Lambda \in \Re^{g \times g}$, i.e. which contains information concerning the annualised average rate of emigrating from any credit-rating grade to another thus:

$$\langle \Lambda \rangle_{ij} = \begin{cases} Transition\,Intensity\,to\,Grade\,i\ \ from\,Grade\,j & i \neq j \\ -\sum_{\substack{k=AAA, \\ k \neq i}}^{Default} \langle \Lambda \rangle_{ik} & i = j \end{cases} \tag{10}$$

With this one generator matrix it would be possible to derive transition probability matrix over any time horizon $T - t > 0$ (that is from $t \geq 0$ to $T > t$, in years) using the matrix exponential formula:

$$\Omega_{(T-t)} = e^{\Lambda(T-t)} \equiv I + (T-t)\Lambda + \tfrac{1}{2!}(T-t)^2 \Lambda^2 + \tfrac{1}{3!}(T-t)^3 \Lambda^3 + \cdots \tag{11}$$

As such Λ is said to *generate* Ω. Indeed, Λ is not required to staty constant. For instance, it could itself be random, such as in a Stochastic Transition Intensity Model (Lando, 1998) which also stipulates that the eigenvalues of Λ_t be a linea function of yet another stochastic process $\{Y_t, t \geq 0\}$, hence:

$$\Lambda_t \xrightarrow{\text{Eigenvalue Decomposition}} \Lambda_t = VD_tV^{-1}, \quad \langle D_t \rangle_{ij} = \begin{cases} \alpha_i + \beta_i Y_t & i = j \\ 0 & i \neq j \end{cases} \tag{12}$$

2.3 Default Process Models

Evidently as a simple default event corresponds to Bernoulli random variable, its complete description requires only one statistical parameter $p \in (0,1)$, known specifically as PD within the default risk context. So with sufficient historical default/non-default data for *each group* of credit assets, estimating PD should be a straightforward case of dividing the number of defaulted cases by the total number of cases, say, within a year. This would provide statistically consistent *point estimate* for the underlying PD.

One of the main problems faced in actual application involves so-called low-default portfolio, such as a portfolio of AAA-rated bonds, none of which, throughout history, has yet to default within a 1-year horizon. Another involves estimating PD for credit items with highly individual characteristics, not amenable to classification to any class of credit assets for which PD has been estimated beforehand. On the other hand, there might be other sources of information that could be brought to bear on the statistical analysis of creditworthiness and business potentials of a particular obligor. Or there could be modelling methodologies that better reflect the default process in some details, as opposed to having just the annual PD as the summary statistics. At any rate the idea is to begin by modelling the default process as realistically as possible, from which the PD estimate could then be derived whenever desired. Principally there are 2 groups of such approaches, based on how the default process is to be modelled:

2.3.1 Structural Default (Asset-value) Models – begins with identifying equity position's limited liability as a *long call option* in firm assets, whereupon the application of Black-Scholes-Merton (Black & Scholes, 1973; Merton, 1973) options pricing theory yields, in addition to the modeled call price, the risk neutral probability by which the call option will expire 'in the money', this probability thus equates to '1 - PD' estimation for default risk application. The model specifies the Asset Value Process $\{S_t, t \geq 0\}$ of an obligor, using the well-known Geometric Brownian Motion (GBM) (Karatzas & Shreve, 1991):

$$\frac{dS_t}{S_t} = \mu \cdot dt + \sigma \cdot dW_t \ , \quad dW_t \sim \mathrm{N}(0, dt) \qquad (13)$$

Further assume, for the sake of simplicity, that the rhs of the obligor's balance sheet consists of 2 items, namely the equity capital C_t and a single liability in the form of a bullet bond with remaining maturity $T - t$ and face value K, so that at maturity the equity will be valued:

$$C_t = \max\{0, S_t - K\} \qquad (14)$$

Should the default event is defined simply as:

$$D = \begin{cases} 1 & S_T < K \\ 0 & otw. \end{cases} \qquad (15)$$

Then at any time $t < T$, it would be possible to estimate the value of the PD parameter (approximately, since real investors do not make decisions under 'risk neutrality') from the Black & Scholes' Call Pricing Formula, which specifies Call Delta thus:

$$D \sim Bernoulli(p)\,,$$
$$where\ p \cong 1 - \underbrace{\mathrm{N}(d_1)}_{\text{"Call Delta"}} \ , \quad d_1 \equiv \frac{\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T - t)}{\sigma\sqrt{T - t}} \qquad (16)$$

But should default event be defined thus:

$$D = \begin{cases} 1 & \exists\, t \in (0, T] \ni S_t < \theta \ , \quad \theta > K \\ 0 & otw. \end{cases} \qquad (17)$$

such as when the bond covenant contains some sort of default acceleration clause over and above simply $S_T < K$, then the situation would resemble pricing a Knock-out Call which loses all its value (the call option is cancelled, with no outstanding obligation left) whenever and as soon as the asset value falls below $\theta > K$ within the period $0 < t \leq T$. Such is referred to as (Black & Cox's) 1st passage approach (Black & Cox, 1976).

At any rate, in actual applications the point is hardly ever about calculating d_1 (whose calculation requires not just the marked-to-market value of firm assets, but also an accurate estimation of the σ parameter governing the asset value process), but the true contribution of this model is to iterate the concept that PD is a function of a 'distance' between asset and liability, weighted in units of asset volatility σ, whereby such a statistical quantity based on $(S_t - K)/\sigma$ has been termed the distance-to-default (D2D) (Avellaneda & Zhu, 2001). For example, one may perform a sort of statistical calibration between ex ante D2D and ex post default event in order to come up with a statistically verifiable mapping between D2D and PD estimate. Such a data calibration-based mapping can then be encapsulated and sold as a commercial software

2.3.2 Default Intensity (Reduced-form) Models – begins with identifying corporate default process with human mortality, production faults, etc., which focuses expediently on the rate by which infrequent events arrive, known in default risk application context as the (Forward) Default Intensity/Hazard Rate $\lambda > 0$ (Duffie & Singleton, 2003), where an interval of time over which the default intensity is relatively high will correspond to a relatively high PD value. This default intensity may be a simple numerical constant, a means parameter governing the distribution of some random variable, a function of time or other variables, or itself a random variable. Indeed the latter case, known as (Doubly) Stochastic Default (Intensity) Models, may borrow heavily from various elements within the Term Structure of Interests Modelling framework build around the instantaneous forward/short rate (evolution of an idealised rate of interest over an instantaneous moment in time) concept, for example, in using the Cox-Ingersoll-Ross (CIR) (Cox, Ingersoll, and Ross, 1985) type of stochastic process to drive stochastic default intensity thus:

$$d\lambda_t = \kappa(\theta - \lambda_t) + \sigma\sqrt{\lambda_t}\,dW_t \ , \quad dW_t \sim \mathrm{N}(0, dt) \qquad (18)$$

2.4 Credit Portfolio Models

2.4.1 Default/Rating Transition Correlation Approaches – for a credit portfolio, i.e. simultaneous, partial exposures to credit risks of $n > 1$ obligors, begin, as per single-obligor case, to define individual default events with Bernoulli random variables, when the objective of the exercise is to find the total number of default events over a period of time for a particular credit portfolio:

$$D_i \sim Bernoulli(p_i), \quad i = 1, \ldots, n$$
$$D_{Portfolio} \equiv \sum_{i=1}^{n} D_i \tag{19}$$

As such, the statistical description of $D_{Portfolio}$ requires 2 elements, namely 1. the individual PD parameters p_i , $i = 1, \ldots, n$, and 2. the correlation structure that exists amongst the random variables D_i , $i = 1, \ldots, n$ themselves. Initially, consider the following special cases

- The case with Homogeneous + Perfect Default Correlation: For a homogeneous portfolio, specify that every obligor shares the same PD value, $PD_1 = \cdots = PD_i = \cdots = PD_n \equiv p$ (In practice, banks may even use PD values to partition a loan portfolio into sub-portfolios). Then hypothetically if every obligor has perfect default correlation with one another, sharing a common 'fortune', the result is simple, but also useless as it is completely unrealistic:

$$D_{Portfolio}^{PerfectCorrel} = nD, \quad D \sim Bernoulli(p) \tag{20}$$

- The case with Homogeneous + Zero Default Correlation: Again start with a homogeneous portfolio, but instead proceed to assign zero default correlations among all the obligors, whence implying default events are independent identically distributed random variables (i.i.d.r.v.). This results in a reasonable model for uses with 'fully diversified' retail loan portfolios, with the total number of defaults $D_{Portfolio}$ distributed according to the Binomial distribution:

$$D_i \sim Bernoulli(p_i), \quad i = 1, \ldots, n$$
$$D_{Portfolio}^{Independent} \equiv \sum_{i=1}^{n} D_i \sim Bin(n, p) \tag{21}$$

- The case with Homogeneous + Zero Rating Correlation: For situations involving obligors with agency credit-rating grades (assuming there are $g > 2$ grades in all), should the i.i.d.r.v. property be preserved, then the result would be similar to above, except there will be instead a credit-rating grade vector distributed according to the Multinomial distribution:

$$\forall i \in \{1, \ldots, n\}, \quad \begin{cases} \Pr(G_i = 1) = p_1 \\ \vdots \\ \Pr(G_g = 1) = p_g \end{cases}$$
$$G_{Portfolio} \equiv \begin{Bmatrix} \sum_{i=1}^{n} If(G_i == 1) \\ \vdots \\ \sum_{i=1}^{n} If(G_i == 1) \end{Bmatrix} \sim Multinomial(n, p_1, \ldots, p_g) \tag{22}$$

The case with Heterogeneous + Zero Rating Correlation: Instead of working with a rating probability vector for a given credit-risky bond, one can work with a joint rating probability matrix for $n = 2$ credit-risky bonds (but no more, as a matrix is a 2-dimensional object, beyond which the mathematical handle comes in the form of n-dimensional tensor). With zero correlation, such a matrix could be obtained by straightforward performing an 'outer' multiplication between the two rating probability vectors. Consider, for example, a portfolio of one 'BB' bond and one 'A' bond, the difference accounts for why such a combination is said to form a heterogeneous portfolio (strictly pedagogical example, as no real-life bond portfolio is likely to be of size $n = 2$):

$$P_{A,BB}^{(1)} = P_{BB,A}^{(1)} = P_{BB}^{(1)} \left(P_A^{(1)}\right)^T = \Omega x_{BB} (\Omega x_A)^T = \Omega x_{BB} x_A^T \Omega^T \tag{23}$$

In reality, however, default correlation will neither be perfect (default correlation of one) nor totally absent (zero default correlation), but will be something in-between the two polar opposites, making portfolio-level analysis of credit risk rather difficult. This is in stark contrast to modelling return correlations among equity shares, where individual returns are generally normally distributed $R_i \sim N(\mu_i, \sigma_i^2), \quad i = 1, \ldots, n$ to begin with, and with well-defined, statistically estimable ***correlation*** parameters, it is straightforward to generalise into $n > 1$ dimensional returns thus:

$$R \equiv \begin{pmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_n \end{pmatrix} \sim \mathrm{N}(\mu, \Sigma), \quad \langle \Sigma \rangle_{ij} = \begin{cases} \sigma_i^2 & i = j \\ \sigma_{ij} = \rho_{ij}\sigma_i\sigma_j & i \neq j \end{cases} \tag{24}$$

Nonetheless methods for analysing correlated credit risks do exist; they come in 3 flavours:

- Homogeneous Bernoulli Mixture – attempts to achieve realism that lies between perfect and zero default correlation. This it does by enabling the PD parameter to change randomly, on the one hand, and letting individual default events be i.i.d., on the other, but only once the value of the PD parameter is determined. So individual obligor defaults independently, all the while utilising the same stochastic parameter, that perhaps reflects different stages of the business cycle, i.e. low PD for cycle peaks, and so on. Since the distribution of the sums of i.i.d. Bernoulli random variables is in fact the Binomial distribution, the result is equivalent to taking a number of binomial distributions and 'mixing' them up in the following manner:

$$\begin{aligned} & D_i \sim Bernoulli(P), \quad i = 1,\ldots,n \\ & P \in \{p_1,\ldots,p_k\}, \quad \Pr(P = p_i) = \theta_i, \quad i = 1,\ldots,k \\ & \quad \Rightarrow \left(D_{Portfolio}^{BernoulliMixture} \mid P = \rho\right) \sim Bin(n,\rho) \\ & \quad \Rightarrow p(x) \equiv \Pr\left(\left(D_{Portfolio}^{BernoulliMixture} \mid p\right) = x\right) = \tfrac{n!}{x!(n-x)!}\rho^x(1-\rho)^{n-x} \end{aligned} \tag{25}$$

For example:

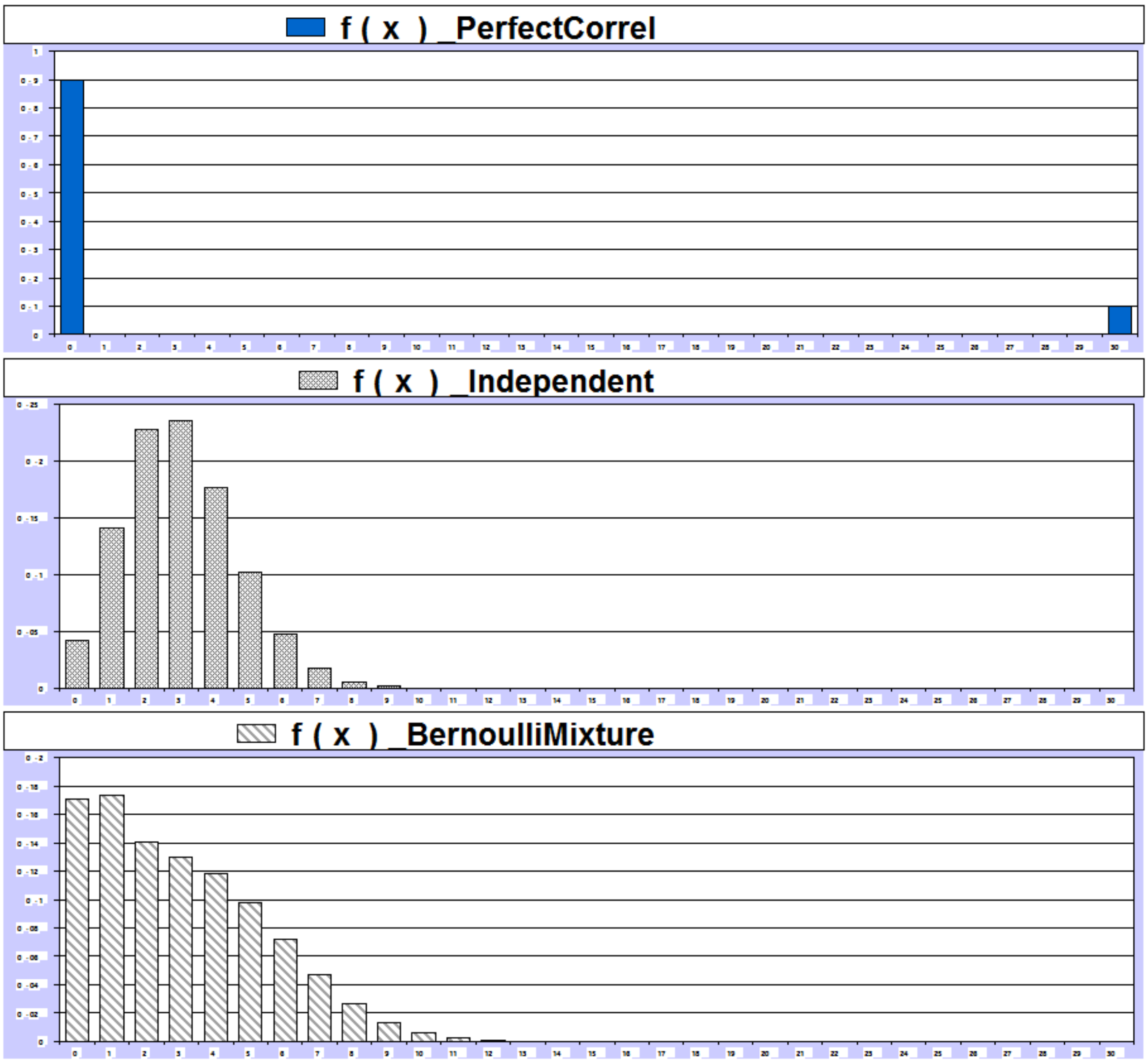
$$\left.\begin{aligned} & n = 30 \\ & P \in \{2.5\%, 10\%, 17.5\%\}, \\ & \quad \theta_1 = \theta_2 = \theta_3 = 1/3 \\ & \therefore (p_1 \quad p_2 \quad p_3)\begin{pmatrix}\theta_1 \\ \theta_2 \\ \theta_3\end{pmatrix} = 0.1 \end{aligned}\right\} \Rightarrow \begin{cases} \mathrm{E}\left[D_{Portfolio}\right] = 0.3 \\ StdDev\left(D_{Portfolio}\right) = 2.44 \end{cases} \tag{26}$$

Note that by comparing with the extreme (perfect or zero correlation) cases at the same mean:

$$\left.\begin{aligned} n = 30 \\ p = 0.1 \end{aligned}\right\} \Rightarrow \begin{cases} \mathrm{E}\left[D_{Portfolio}^{PerfectCorrel}\right] = \mathrm{E}\left[D_{Portfolio}^{Independent}\right] = 0.3 \\ StdDev\left(D_{Portfolio}^{PerfectCorrel}\right) = 9 \\ StdDev\left(D_{Portfolio}^{Independent}\right) = 1.64 \end{cases} \tag{27}$$

It is clear how the Binomial Mixture model yields a standard deviation which is somewhere between the two extreme cases. Shape-wise, the Bernoulli mixture distribution is also somewhere between the bi-valued shape of the Bernoulli and the hill-like shape of the binomial:

Figure 2. Portfolio Default Correlation – Perfect Correlation vs. Independent vs. Bernoulli Mixture



Latent-Variate Joint Distribution (details and example calculation for the case of $n = 2$ obligors to be found in (Crouhy, Galai, and Mark, 2000) – begins with the observation that generalizing from *univariate* to *multivariate Normal* simply involves introducing the correlation parameters to within the variance-covariance matrix, but an extension from univariate Bernoulli to "multivariate Bernoulli" distribution is not available. So the approach here is to convert univariate Bernoulli distribution into univariate Normal distribution first, extend to multivariate Normal distribution by incorporating the correlation parameters, then convert back to a sort of "multivariate Bernoulli" distribution, furnished with a hidden correlation structure. In practice it's arguably acceptable to proxy this type of credit-risk related correlation structure with market-risk related structure of equity-return correlations, citing the same sort of arguments used by structural default models to link market risk factors with credit risk factors:

$$A \equiv \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_n \end{pmatrix} \sim \mathrm{N}(\mathbf{0}, \Theta), \quad \langle \Theta \rangle_{ij} = \begin{cases} 1 & i = j \\ \sigma_{ij} = \rho_{ij} & i \neq j \end{cases} \tag{28}$$

$$\Rightarrow \begin{cases} p_i \equiv \Pr(D_i = 1) = \Pr(A_i \leq \theta_i), \quad \theta_i = \mathrm{N}^{-1}(p_i) \\ \Pr(D_i = 1 \cap D_j = 1) = \Pr(A_i \leq \theta_i \cap A_j \leq \theta_j) \end{cases}$$

What is good about this is that extending from multi-obligor *default* to multi-obligor *rating* correlation is straightforward. The downside is that in practice, it's impossible to calculate directly the joint default/rating probability for a portfolio of size $n > 2$

Distributional Concordance Copula (Nelson, 1999) – Whereas Homogeneous Bernoulli Mixture and Latent-Variate Joint Distribution introduce correlation-inducing parameters in the beginning of the analysis, Distributional Concordance Copula cleverly allows the obligors' individual probability distributions to be integrated together at the end of the analysis:

$$C_{ij} : [0,1]^2 \to [0,1]$$
$$\ni \Pr(A_i \le \theta_i \cap A_j \le \theta_j) = C_{ij}\left(\Pr(A_i \le \theta_i), \Pr(A_j \le \theta_j)\right) \quad (29)$$

As normally these types of models will require complex computations to obtain numerical result for which there is often no closed-form mathematical expression, the use of Numerical Methods (Brandimarte, 2001) and/or Monte Carlo Simulation (Glasserman, 2004) will be necessary.

2.4.2 Stochastic Loss Arrival/Convolution Approaches – involves 2 modelling steps: modelling (Default) Count and (Default) Severity, followed by synthesising the two into the total loss distribution the aggregate, whereby individual default severities, one from each default event, are added up to register a single value, the essential feature here being that the number of defaults is given by default count, a random quantity.

For default counts, start with time $t = 0$, then 'wait' for the moment any obligor in said credit portfolio defaults first. Define the interval taken for this to occur as the first random variable $\tau_1 > 0$. Then define the time interval between this and the moment of the second default event as the second random variable $\tau_2 > 0$, and so on, whence the series forms the stochastic process interarrival times $\{\tau_1 > 0, \tau_2 > 0, \ldots\}$, where each *individual* intervals are generally *assumed* to be independently distributed of one another, i.e. $\forall\, i, j \in \{1,2,\ldots\}, i \ne j$, $\mathrm{E}[\tau_i \tau_j] = \mathrm{E}[\tau_i]\mathrm{E}[\tau_j]$. The default count over $(0, t]$ can simply be defined in terms of the following counting process:

$$N(t) = \begin{cases} 0 & \tau_1 > t \\ 1 & \tau_1 = t \\ \sup_n \left\{\sum_{i=1}^{n} \tau_i \le t\right\} & otw. \end{cases} \quad (30)$$

As for default severity, given that default count turns out to be $N(t) = n$, $n \in \{0,1,\ldots\}$: if $n = 0$, then there is no default severity to consider; if $n = 1$, then there is one random variable corresponding to the singular default severity; and if $n > 1$, then there would be a stochastic process depicting a *series* of default severities $\{L_1 > 0, L_2 > 0, \ldots, L_n > 0\}$, where each *individual* loss is generally *assumed* to be independently distributed, i.e. $\forall\, i, j \in \{0,1,\ldots,n\}, i \ne j$, $\mathrm{E}[L_i L_j] = \mathrm{E}[L_i]\mathrm{E}[L_j]$. Canonically:

Compound Poisson Process Model – corresponds to the well-known and widely-used "classic" model in non-life insurance, namely the Cramér-Lindberg Model (Mikosch, 2004), which stipulates the interarrival times $\{\tau_i, i = 1,,2,\ldots\}$ to be i.i.d.r.v. which are distributed according to the following Exponential distribution:

$$i.i.d.r.v.\ \tau_i \sim Exp(\lambda) \Rightarrow \begin{cases} f(\tau_i) = \tau e^{-\lambda\tau} \\ \mathrm{E}[\tau_i] = \dfrac{1}{\lambda} \end{cases}, \quad i = 1,,2,\ldots \quad (31)$$

This is equivalent to saying that default arrivals follows a Poisson Process, so that the default count is consequently distributed according to the following Poisson distribution:

$$N(t) \sim Poisson(\lambda t)$$
$$\Rightarrow \begin{cases} \Pr(N(t) = x) \equiv p(x) = \dfrac{e^{-\lambda t}(\lambda t)^x}{x!}, \quad x = 0,1,2,\ldots \\ \mathrm{E}[N(t)] = \lambda t = Var(N(t)) \end{cases} \quad (32)$$

For example, by letting default severities $\{L_k, k = 1,,2,\ldots,n\}$ be i.i.d.r.v. with Gamma distribution:

$$L_{Portfolio}(t) = \sum_{k=1}^{N(t)} X_k \ , \quad N(t) \sim Poisson(\lambda t) \ , \quad i.i.d.r.v.\ X_k \sim Gamma(\alpha, \beta) \ , \quad k = 1,\ldots,N(t) \tag{33}$$

As a second example, suppose that in the event of default, there is a probability $1-\rho > 0$ of *full recovery*, and a probability $\rho > 0$ of partial recovery at fixed $\kappa > 0$ default severity, then the total loss would be $\kappa > 0$ multiple of Bernoulli random variables thus:

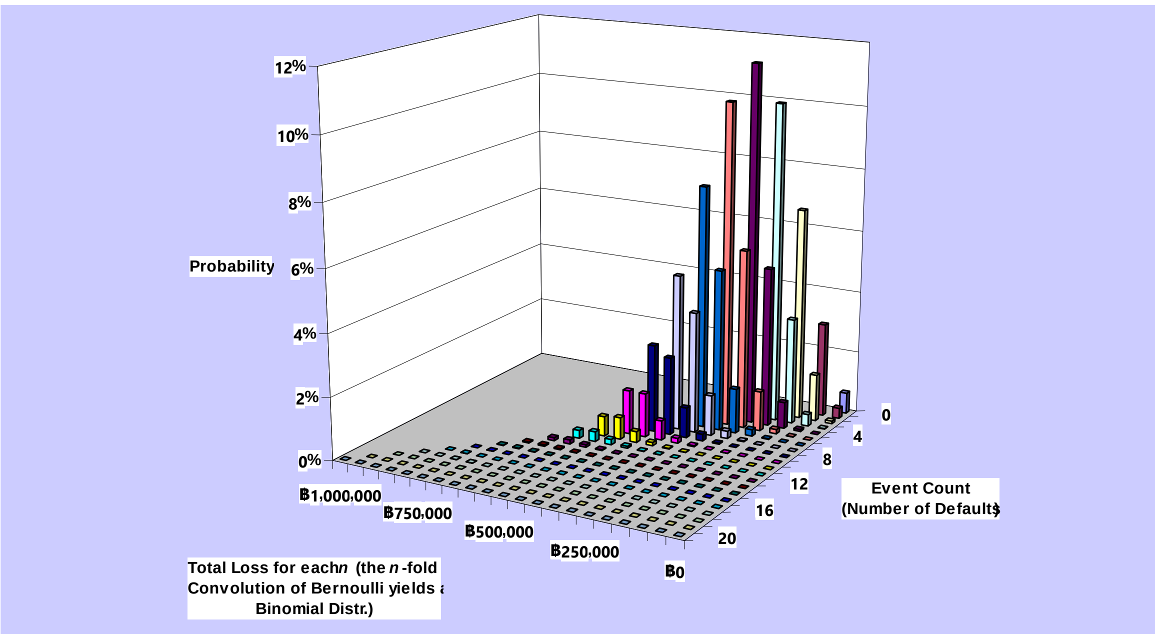
$$L_{Portfolio}(t) = \sum_{k=1}^{N(t)} X_k \ , \quad N(t) \sim Poisson(\lambda t) \ , \quad i.i.d.r.v. \left(\frac{X_k}{\kappa}\right) \sim Bernoulli(\rho) \ , \quad k = 1,\ldots,N(t) \tag{34}$$

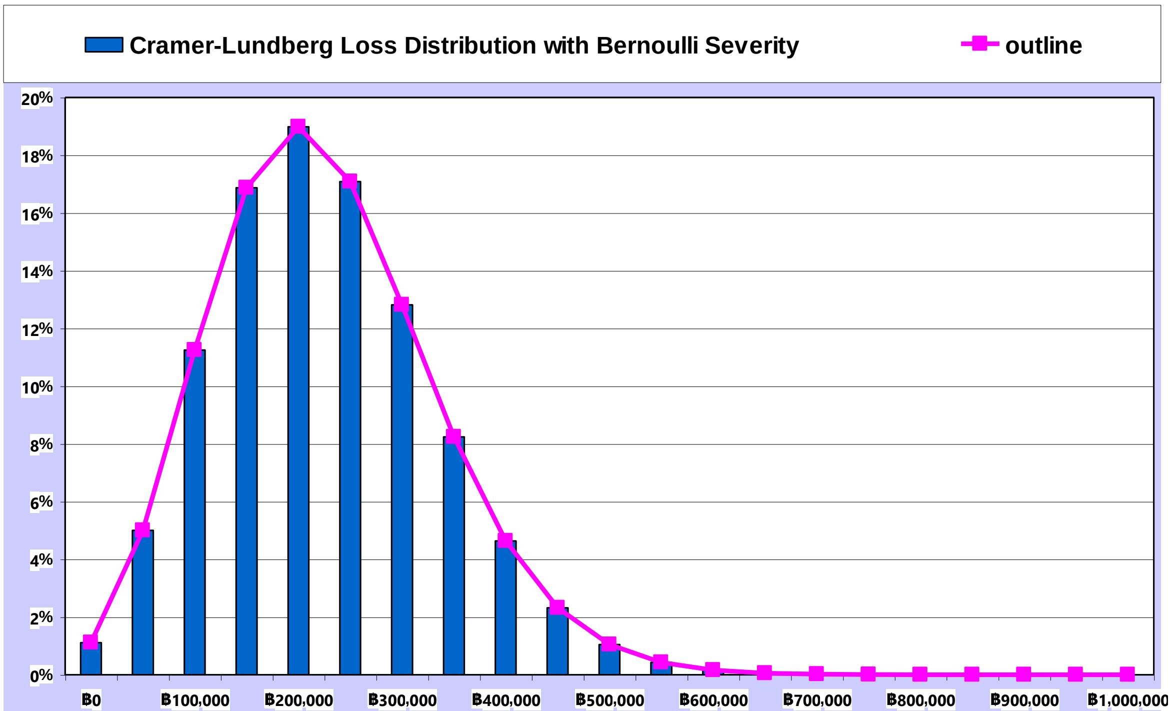
But of course whenever it's determined that $N(t) = n$, such a sum of Bernoulli i.i.d.r.v. would be distributed according to a simple Binomial distribution hence:

$$\begin{aligned} & N(t) \sim Poisson(\lambda t) \ , \\ & \frac{1}{\kappa}\left(L_{Portfolio}(t) \,\middle|\, N(t) = n\right) \sim Bin(n, \rho) \ , \\ & \therefore\ p\left(\frac{l}{\kappa}\right) \equiv \Pr\left(\frac{L_{Portfolio}(t)}{\kappa} = \frac{l}{\kappa}\right) \\ & = p(n) \cdot p\left(\frac{l}{\kappa} \,\middle|\, n\right) = \left(\frac{e^{-\lambda t}(\lambda t)^n}{n!}\right) \cdot \left(\frac{n!}{\left(\frac{l}{\kappa}\right)!\left(n-\frac{l}{\kappa}\right)!}\rho^{\frac{l}{\kappa}}(1-\rho)^{n-\frac{l}{\kappa}}\right) \end{aligned} \tag{35}$$

For a numerical example, let $\{\lambda t = 5; \rho = 0.9; \kappa = THB50{,}000\}$:

Figure 3. Cramer-Lundberg Total Loss Distribution with i.i.d.r.v. Bernoulli Individual Loss Severity





Note how this ends up essentially with $L_{Portfolio}(t) = \kappa L$, $L \sim Poisson(\lambda^{\circ})$, $\lambda^{\circ} \cong 4.5$. In other words, specifying $\rho = 0.1$ as the probability of full recovery equates to reducing the default arrival rate from $\lambda t = 5$ (with no possibility of a full recovery) to $\lambda^{\circ} t \cong \underbrace{(1-\rho)}_{0.9}\underbrace{\lambda t}_{5} = 4.5$, as theory would have it. (Johnson, Kotz, and Balakrishnan, 1994; Meintanis & Koutrouvelis, 1999). So theoretically this example is strictly demonstrative, as it represents no advance in modelling realism.

Renewal Model (Embrechts, Klüppelberg, and Mikosch, 1997) – slightly expands the scope of applications from the compound Poisson process model by enabling interarrival times $\{\tau_i, i = 1,,2,\ldots\}$, still i.i.d.r.v., to be distributed more generally, so long as the means is finite:

$$i.i.d.r.v.\ \tau_i \in \Re^+ = (0,\infty) \ni \mathrm{E}[\tau_i] = \frac{1}{\lambda}, \quad i = 1,2,\ldots \tag{36}$$

3. CONCLUDING REMARKS

It can be said that the ultimate aim of modelling financial risks, be they credit, market/price, or operational, is to establish individual and portfolio-level gain/loss distributions. Quick comparisons reveal that for market/price risks: (i) gain/loss distribution is on the whole symmetric, with nearly as much upside as downside opportunities (ii) this gain/loss distribution always exhibits itself (due to continuous marking-to-market) and (iii) in the aggregate, correlation structure plays the key role in shaping the risk profile at the portfolio level. For operational risks: (i) there is only severity (of loss) distribution, with no upside to speak of, (ii) this severity distribution, by the very nature of 'high impact' risk, materialises once in a long while, and (iii) in the aggregate, correlation structure is nearly meaningless. In this light, credit risks lie somewhere in between: (i) loss distribution is mostly about downside potentials, with occasional upside potentials in the form of credit-rating upgrades, or surprised gains from collateral disposals, (ii) this loss distribution manifests itself regularly, not updating as frequently as for market/price risks, nor as elusive as in the case of operational risks, unless one focuses strictly on *low default portfolio*, and (iii) in the aggregate, correlation structure plays a crucial, but not overwhelming role in the analysis of portfolio-level exposure. Indeed credit correlation structure is crucial as a basis for a thorough process of credit risk analysis, yet while the technology is not so widespread, proxy mechanisms, such as Basel II's IRB RWF, may be forced into service, using little more than information involving just single-obligor loss distribution.

This article surveys and organises credit risk models, specifically the loan discriminant, rating transition, and default process models of single-obligor credit risks, which are most relevant to the capital adequacy framework *a la* Basel II, as well as a number of credit portfolio levels, which henceforth shall form the backbone of credit risk modelling and analysis for financial institutions.

REFERENCES

1. Altman, Edward I. (1968), "Financial Ratios, Discriminant Analysis and the Prediction of

Corporate Bankruptcy", Journal of Finance, September: 589-609.
2. Avellaneda , M. & Zhu, J. (2001), "Modeling the Distance-to-Default Process of a Firm", working paper, available online via http://www.defaultrisk.com/pp_model_57.htm.
3. Black, F. & Cox, J. (1976), "Valuing Corporate Securities: Liabilities: Some Effects of Bond Indenture Provisions", Journal of Finance, 31: 351-367.
4. Black, F. & Scholes, M. (1973), "The Pricing of Options and Corporate Liabilities", Journal of Political Economy, 81 (3): 637-654.
5. Boser, Guyon, and Vapnik (1992), "A training algorithm for optimal margin classifiers", in Haussler, D. (ed.), 5th Annual ACM Workshop on COLT, 144-152, Pittsburgh: ACM Press.
6. Brandimarte, Paolo (2001), "Numerical Methods in Finance: A MATLAB-Based Introduction", Wiley-Interscience.
7. Cox, Ingersoll, and Ross (1985), "A Theory of the Term Structure of Interest Rates", Econometrica, 53: 385-407.
8. Chaplin, Geoff (2005), Credit Derivatives: Risk Management, Trading and Investing, Chichester: John Wiley & Sons.
9. Crouhy, Galai, and Mark (2000), "A Comparative Analysis of Current Credit Risk Models", Journal of Banking & Finance, 24: 59-117.
10. Duda, Hart, and Stork (2000), Pattern Classification (2nd ed.), Wiley Interscience.
11. Duffie, D. & Singleton, K.J. (2003), Credit Risk: Pricing, Measurement, and Management, Princeton: Princeton University Press.
12. Embrechts, Klüppelberg, and Mikosch (1997), Modelling Extremal Events for Insurance and Finance, Berlin: Springer.
13. Fisher, R.A. (1936), "The Use of Multiple Measurements in Taxonomic Problems", Annals of Eugenics, 7: 179-188.
14. Glasserman, Paul (2004), Monte Carlo Methods in Financial Engineering (Applications of Mathematics No. 53), New York: Springer.
15. Johnson, Kotz, Balakrishnan (1994), Continuous Univariate Distributions Volume 1 (2nd ed.), New York: Wiley-Interscience Publication.
16. Karatzas, I. & Shreve S.E. (1991), Brownian Motion and Stochastic Calculus (2nd ed.), New York: Springer.
17. Lando, David (1998), "Cox Processes and Credit-Risky Securities", Review of Derivatives Research, 2: 99-120.
18. Lando, David (2004), Credit Risk Modelling, Princeton: Princeton University Press.
19. McFadden, Daniel L. (1974), Applied Logistic Regression (2nd ed.), New York: Wiley.
20. Meintanis, S.G. & Koutrouvelis, I.A. (1999). "Chi-Squared Tests of Fit for Generalized Poisson Distributions Based on the Moment Generating Function", InterStat, May Issue, 1-19.
21. Meissner, Gunther (2005), Credit Derivatives: Application, Pricing, and Risk Management, Malden, MA: Blackwell Publishing.
22. Merton, Robert C. (1973), "Theory of Rational Option Pricing", Bell Journal of Economics and Management Science, 4 (1): 141-183.
23. Mikosch, Thomas (2004), Non-Life Insurance Mathematics: An Introduction with Stochastic Processes (Universitext), Berlin: Springer.
24. Nelson, Roger B. (1999), An Introduction to Copulas (Lecture Notes in Statistics No. 139), New York: Springer.
25. Pao, Yoh-Han (1989), Adaptive Pattern Recognition and Neural Networks, New York: Addison-Wesley.