CERTIFIED EXPERT IN RISK MANAGEMENT

UNIT 4.1: CREDIT RISK MANAGEMENT

AACSB ACCREDITED

EFMD EQUIS ACCREDITED

Frankfurt School
Development Finance | e-Campus

Certified Expert in Risk Management

Unit 4.1: Credit Risk Management

Symbols

Definition

Further Reading

!

Key Message

Example

Exercise

Video

2. edition 09/2018

Content

Abbreviations

ABC	Activity Based Costing
ADB	Asian Development Bank
ALCO	Asset and Liability Management Committee
ALM	Asset-Liability Management
AML/CFT	Anti-Money Laundering / Countering the Financing of Terrorism
ATM	Automated Teller Machine
BCBS	Basel Committee on Banking Supervision
BSC	Balanced Scorecard
CAR	Capital Adequacy Ratio
CGAP	Microfinance Secretariat at the World Bank
COSO	Committee of Sponsoring Organizations of the Treadway Commission
EAD	Exposure at Default
EL	Expected Loss
ERM	Enterprise Risk Management
EUR	Euro currency
FRM	Financial Risk Manager - professional designation
HR	Human Resources
ISO	International Standards Organization
KPI	Key Performance Indicator
KRI	Key Risk Indicator
LGD	Loss given Default
MFI	Microfinance Institution
MIS	Management Information System
MIV	Microfinance Investment Vehicle
MSME	Micro-, Small and Medium Enterprise
NBFI	Non-Bank Financial Institution
NGO	Non-Governmental Organization
NPL	Non-Performing Loan
PAR	Portfolio-at-risk
PD	Probability of Default
PMI	Project Management Institute

PMI-RMP	PMI Risk Management Professional designation
PRM	Professional Risk Manager professional designation
ROA	Return On Assets
ROE	Return On Equity
SME	Small and Medium Enterprise
TA	Technical Assistance
USD	US Dollar
VPN	Virtual Private Network

Learning Outcomes

This Unit 4.1 on credit risk is the single largest unit in the Certified Expert in Risk Management course. It goes to the heart of the matter of risk in financial institutions. The learning objectives are ambitious: We want to teach you everything we know about identifying, measuring, reporting, mitigating and managing credit risk with a special focus on microcredit and SME lending. Once you have worked through the script and all of the exercises, you should be able to:

- Communicate effectively about the parameters that determine the loss distribution in credit portfolios.
- Understand the nature and drivers of default in MSME credit.
- Assemble and interpret descriptive portfolio performance statistics such as arrears schedules, vintage curves or a transition matrix.
- Derive appropriate loan loss provisioning methods in compliance with prudential norms and IFRS.
- Design collateral strategies for various markets and loan products that are both socially responsible and conducive to minimizing loss given default.
- Organize an efficient arrears management and collections process in compliance with responsible finance practices.
- Build and maintain statistical credit scoring and rating models for microenterprise credit and SMEs.
- Incorporate forward visibility of default probabilities into a risk-based credit pricing framework.

1 Introduction and Overview

Now, we are done with the terminological warm-ups and ready to dig into the details, risk-by-risk. This Unit 4.1 and each of the following Units will take one of the core risk dimensions and walk through the full risk management process from identification to measurement to action. Credit risk is up first. Then comes operational risk, interest rate risk, exchange rate risk, and finally we will deal with the main downstream risk category of liquidity.

Unsurprisingly, this credit risk Unit will be the single biggest piece in the course with the greatest number of analytical tools, the most homework and some of the most interesting number crunching.

We will talk about credit transaction versus credit portfolio risk and about organizational principles of credit risk management in SME lending, micro-enterprise finance and consumer credit. Portfolio risk management always starts with a keen eye on concentrations and the need for effective ex-ante diversification and macro-budgeting of exposures by industry and geography. We will also study traditional portfolio performance diagnostics: such as arrears aging schedules, vintage curves and the transition matrix.

Much time will be spent on the data requirements for predictive credit modeling and the development of a comprehensive client data strategy. This data platform will at the same time also enable targeted marketing and credible reporting on the social and economic development impact of financial access.

Assuming that we have available good socio-demographic, financial and credit history data on our clients, we can build statistical models for the probability-of-default and loss-given-default parameters. In fact, you can do this at home with some inexpensive plug-in software for Excel. We will show you step by step how it is done on some examples of real loan portfolio data. The same analytical apparatus will then be used to develop a behavioral scoring model, such as a collections scoring, for example. A collections score will help you decide which clients in arrears might be most responsive to which type of arrears management actions. We will also discuss a detailed borrower and facility rating model for SMEs in emerging and developing markets. And finally, with good estimates for the basic portfolio risk parameters in place, we can now put it all together in a risk-based pricing model.

So let's get started. We repeat the simple definition of credit risk that we introduced in Unit 3:

> ⇨ ***Credit risk*** *is defined as the possibility that a borrower or other contractual counterparty might default, i.e. might fail to honor their contractual obligations*

Let's keep the **counterparty credit risk** element from the above definition on our check list for later, when we will discuss liquidity risk in Unit 4.5. The most likely context in which an MSME bank would encounter the risk of a wholesale finance counterparty defaulting is in treasury management: investing the liquid asset reserve in bank placements and high grade debt securities issued by government, banks and prime corporates.

Counterparty credit risk aside, we should still amend the above definition by a further dimension called **migration risk**. Of course, credit risk is essentially about the loss that occurs, if a borrower does not pay. Yet, credit losses may arise well before a borrower actually misses a payment. Losses can be triggered simply by the fact that the perceived likelihood of a future default has increased while an exposure is outstanding.

> ⇨ *The potential deterioration of the credit quality of an un-defaulted exposure is called* ***migration risk****. This form of potential loss is generally also subsumed under a broader definition of* ***credit risk****.*

Migration risk is not as abstract as it sounds and even has relevance for MSME credit. Just imagine you have a portfolio of microenterprise loans outstanding in an area where the underlying source of economic activity is a big mining site. Tomorrow, there is an accident at the mine. The main shaft floods, production will be idled for at least six months and most laborers at the mine will be laid off. Nothing has changed in your microcredit portfolio for this month's collections, defaults are low. But clearly you can see the train wreck coming. If you happen to routinely put microcredit receivables into a security pool for multiple lenders or even sell loans into a securitization vehicle, these future credit losses will be immediately monetized. A reasonable counterparty would value this loan portfolio much lower on the day after the mining accident. Thus a bigger valuation haircut will be imposed as you borrow against the loans, and selling them might become entirely impossible.

We will come back to the idea of **migration risk** as we consider the pricing of longer-term loans in Chapter 10 of this Unit. If we make loans for several years, we only get one chance up front to assess the risk of default and price it into the loan. Much may happen that would increase the risk of default over time. Thus, we should anticipate a certain downward migration of the credit quality and factor in an additional charge for credit migration risk.

Naturally, from the perspective of a microfinance institution or a traditional retail / SME bank, the big deal when it comes to credit risk is the short-term **borrower default component.** The "simple" question of whether or not the borrower will default over the next year and how much we stand to lose, if he does, is what we will look at under the microscope from all different angles now.

As we zoom in on the risk of small businesses and individuals failing to meet their loan obligations, we will continue to use the distinction between the transaction risk and portfolio credit risk dimensions:

⇨ ***Transaction risk*** *refers to individual loans and essentially measures (1) the standalone probability that the borrower will be able to repay, as well as (2) the ultimate loss in the case of a borrower default after use of collateral and other mitigating factors.*

Portfolio credit risk *is concerned with measuring correlations between individual borrower defaults, the effects of diversification, the cyclicality of collateral values and the implications of reputation and contagion effects in microcredit.*

There is so much to say, so many stories to share, so many models to discuss, when it comes to the default risk of micro- and SME borrowers. Where do we start and how do we bring some conceptual order to all of this?

We propose to use one simple formula as our guiding light through the entire Unit. You have probably seen this before - it has been made famous by the Basel II rules on regulatory capital for credit risk.[1] The financial crisis has dented the confidence in credit risk modeling a bit. But nonetheless, this basic little formula was not invented by the Basel Committee and it is also not discredited by some of the shenanigans that banks have used to keep their capital requirements low. It is as simple and always true as saying the 'world is round':

[1] See paragraphs 211ff in BCBS 2004 Basel II (June 2006 compilation).

Expected Loss = Probability of Default * Exposure at Default * Loss Given Default

Or, for short we will write: **EL=PD*EAD*LGD.**

By itself this "model" does not explain anything, of course. It is just a very simple way to structure our thinking about the losses from borrower default into three elementary dimensions that we can take apart in more detail and then develop specific models for each. So, instead of just saying "credit risk is when the guy does not pay the loan back", we look at this loss as arising from three distinct factors:

1) There must be a **default**. This is an event or behavior which is a characteristic of the borrower.
2) The loss then depends on how much the client owed the particular lender when he stopped paying, i.e. the **exposure at default**.
3) And finally, we need to factor in **how much of that exposure will actually be lost** after we liquidate any collateral and attempt to collect through the legal and arrears management routes.

2 The EL=PD*EAD*LGD Way of Thinking

Let us unpack **the EL=PD*EAD*LGD** logic further. For this, we first need some more specific definitions:

➯ *We define the **Expected Loss (EL)** as the average or mean amount of credit loss to be incurred over a particular time period. The loss is measured as the present value or book value of receivables that will not be collected or will have become unrecoverable and therefore will be written off or otherwise expensed during a particular period of time.*

*The **Probability of Default (PD)** is the percentage probability of a borrower entity to produce a default event as perceived by the lender over a specified period of time, typically one year. The PD is most often stated for a future period beginning immediately, but can also be expressed as a forward default probability beginning in one year for one year, for example.*

*The **Exposure at Default (EAD)** is the total balance owed by the borrower to the particular lender at time of default expressed in currency units.*

***Loss-given-default (LGD)** is the percentage of EAD that is considered lost, once it has been established that a default has occurred. The LGD is equal to 100% minus the percentage of EAD that will be recovered by way of liquidation of collateral and other post-default collection actions. For the purposes of establishing LGD, the post-default cash flows from recoveries must be discounted back to the time of default at the original internal rate of return of the defaulted contract.*

Figure 1 illustrates the relationship between the LGD and the net present value of post-default recoveries. IAS 39 and the new IFRS 9 standard require that receivables for which evidence of impairment exists be carried on the balance sheet at the net present value of residual realizable cash flows. The discounting should be done at the effective rate (internal rate of return) of the original loan contract. Obviously, materialized default is a very clear "evidence of impairment", so the impaired receivables valuation under IAS 39 / IFRS 9 applies. We will get back to LGD, collateral, IAS 39 / IFRS 9 and impairments and provisioning in more detail in chapters 5 and 6. We just wanted to give an initial understanding of LGD here, so we can appreciate the EL=PD*EAD*LGD logic.

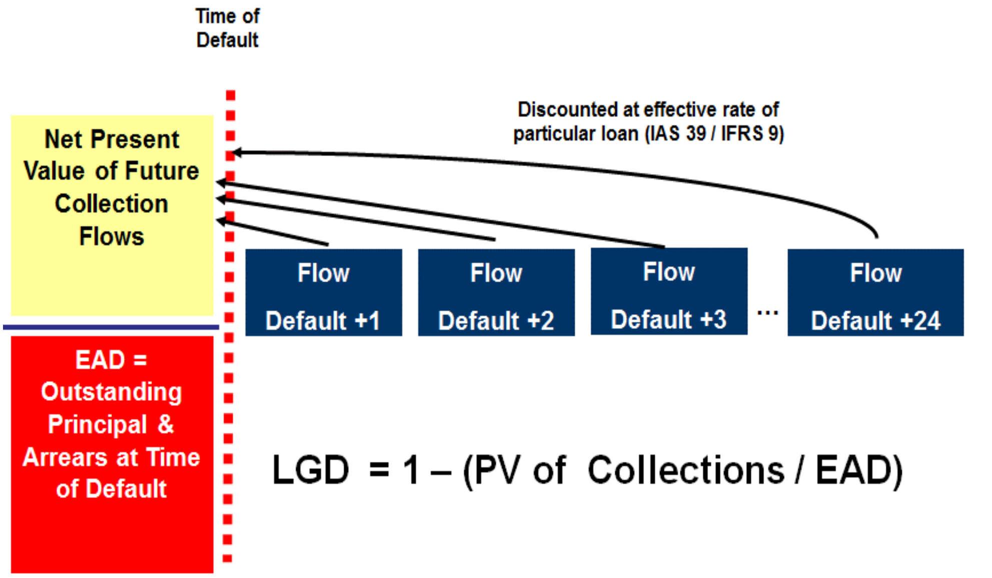

Figure 1: Definition of LGD and Relationship to Present Value of Recovery

We should also expand a bit on the need for a separate **EAD dimension**. After all, is the amount disbursed or the amount outstanding today not the obvious measure of exposure? Absolutely, the amount outstanding today is probably a good starting point for how much gross exposure is at risk, should the particular borrower default. However, the point is that we want to establish an expected loss for a forward period of time, say for a year starting tomorrow. Since the default may occur at any time over the next year, it is obvious that the amount outstanding at the time of default is also an uncertain "random" variable that has some relationship with the amount outstanding today, but by no means must be equal to it.

✓ **Example: EAD**

Let's imagine a simple installment loan that is reimbursed in 12 equal monthly installments annuity style. The loan is disbursed today and we wonder how much might be outstanding, if and when the borrower defaults over the course of the coming year. If we know or assume that the borrower has a PD of 3% for the year, we could surmise that default is equally likely throughout the year and that the borrower would default with 3%/12 every month and that on average, the default would occur just after the sixth installment due date. If it was a loan with linear principal reimbursement, the expected EAD at that time would be 50% of the starting balance. On annuities, the principal would still be higher, of course. Exactly 52.5%, if the loan had a nominal annual rate of 20%.

Exercise M4:1_Ex1: Annuity loan formulas

We can't let this opportunity pass to do another little finger exercise in annuity formulas in Excel: What is the principal balance outstanding at the end of month 5 (just after receipt of the 5th installment) for an annuity style loan of 1,000 that is reimbursable in 18 equal monthly installments and that carries an annual nominal rate of 24% (i.e. 2% per month)?

Hint: remember the annuity formulas =PV(), =PMT(), =PPMT(). Also consider that the principal balance is always equal to the present value, =PV(), of the future loan payments discounted at the loan rate.

Solution: 756.96. See M4.1_Ex1_Annuity.xlsx for details.

Back to the idea that the default might occur on average half way through the year with an exposure of default that has been diminished by six interim principal installments. Sorry, to have led you on. We did learn a new trick in Excel along the way maybe, but this idea of EAD being significantly smaller than the current exposure outstanding just does not hold water. Not even in the simplest microcredit cases, where we only make fixed installment loans for tenors of around one year, one client, one loan at a time. In fact, we venture say that as a reasonable assumption we should go with the notion that **EAD will always be larger than the current un-defaulted balance outstanding** for the particular borrower. Un-defaulted means that there are no pre-existing arrears on this loan. Here is why EAD tends to be larger, not smaller than the balance outstanding today:

1) You could say, if we are only considering currently un-defaulted exposures and default might be commonly defined as reaching 90+ days in arrears, well than nobody can default in the first three months of the observation period, because they could not have reached 90+ days in arrears. And thus the balance outstanding would have been reduced by three principal installments in the meantime. But this argument is not true, because the client could certainly stop paying with the first installment due in the observation period. By the time he reaches default, he would then precisely not have made those three installments. And that means that the balance outstanding when we finally register the default at 90 days has not diminished at all. Instead it increased by the interest not paid with the previous missed installments.

2) When corporations go bankrupt, they always go down in flames with the last dollar on the last credit line drawn to the limit. Think of Enron or Worldcom and other famous disasters. Default is by definition characterized by a shortage of liquidity and the desperate search for fresh credit, until the lenders pull the plug and say: "Not a cent more".

3) The corporate example shows nicely that the balance outstanding profile over the course of the next year, is not independent of the default variable. If the client takes the default bullet from the Russian roulette, so to speak, we already know that the balance outstanding profile will have evolved differently until the time of default, than if all had been going well for the company. Default maximizes the balance outstanding profile. Thus, EAD can always be assumed greater than current exposure. How much greater, naturally depends on institutional practice and lending policies, the prevalence of open limits and overdraft lines etc. It must be studied and modeled empirically.

4) The corporate credit logic of EAD being greater than current exposure actually transfers seamlessly to microcredit. Also in microcredit, a default will be preceded by a deterioration of the borrower's financial situation. There will be attempts to make up for the liquidity shortage by borrowing more: from your institution, from a competitor, from family or from the village loan shark. Now, think of the prevalent practice of early settlement and re-advance in micro-lending. For the most part, these are good borrowers requiring more cash more quickly to grow their business. But if you are not watching it closely, the future defaulters who are already short on cash because their business is failing, might be using the same process to top off their loan balance with the MFI one last time, before they will never pay again. And even though in microcredit, it is not common practice to grant multiple parallel loans or revolving lines of credit, many institutions do offer emergency and other short-term special purpose loans in parallel to the main working capital product. Et voilà, what could be a more obvious emergency than a failing business that needs cash urgently.

5) Finally let's consider the effect of the "graduation principle" in microcredit. This means that borrower relationships are built up gradually over time with ever increasing loan amounts in each loan cycle. So, even if all goes as planned and there is no deterioration of the financial situation of the client, it is quite likely, that over the course of a year, the current loan would have been entirely paid off and replaced by a new larger loan. The new loan would be early in its maturity

and most likely display a higher balance outstanding than we see on the current loan at this time.

Thus all told, even in microcredit, **EAD will typically exceed current exposure** to the particular lender.

Now, that we have a bit of a feeling for the nature of the three components in the EL formula, let's look again at the whole picture of **EL=PD*EAD*LGD.**

EL=PD*EAD*LGD as Random Variables

On the day that we disburse a loan and look ahead at the next year, all three factors (PD, EAD, LGD) are what we call random variables.

> ⇨ *In probability and statistics, a **random variable or stochastic variable** is a variable whose value is subject to variations due to chance. As opposed to other mathematical variables, a random variable conceptually does not have a single, fixed value (even if unknown). Rather, it can take on a set of possible different values, each with an associated probability.*
>
> *A random **variable's possible values** might represent the outcomes of a yet-to-be-performed experiment or an event that has not happened yet, or the potential values of a past event whose already-existing value is not yet known.*
>
> *A random variable can be classified as either **discrete,** i.e. it may assume any of a specified list of exact values, or as **continuous**, i.e. it may assume any numerical value in an interval or collection of intervals.*
>
> *A discrete random variable that can take one of a limited, and usually fixed, number of possible values is called a **categorical variable.** A categorical variable that can assume exactly one of two possible values (e.g. [yes; no] or [0; 1]) is termed separately as a **binary variable** or **dummy variable**.*
>
> *The mathematical function describing the possible values of a random variable and their associated probabilities is known as a **probability distribution.***

The elements PD, EAD and LGD each are random variables for every individual loan. EL as a function of these three random variables then **is also a (derived) random variable.** Its value depends on the outcome of each of the three underlying random variables.

At the same time, we can think of EL and its components as random variables at the aggregate portfolio level.

Just as EL_i (i.e. the EL for an individual loan) is a random variable that is derived from the outcomes of **$PD_i*EAD_i*LGD_i$**, we can look at EL_p as an aggregate random variable **$EL_p=PD_p*EAD_p*LGD_p$** that is the result of the summation of individual loan losses to the portfolio level.

Let's experiment with these random variable concepts a little in Excel. The notion of default is a binary random variable that is most frequently coded like this: no default = 0, default =1. The probabilities of the two outcomes are expressed typically for a specific period of time, often for a year. The specific time horizon is obviously necessary, otherwise the question would resemble the long-run probability of death, which is 100% always.

✓ Imagine a borrower who has a PD of 5% per annum. This means he would default 5 times in 100 years. Or, among 100 identical and independent borrowers each with a 5% PD, in an average year, five would have defaulted and 95 would still be in good credit standing at the end of the year. Yet, in loan default, there is not just one "lottery drawing" at the end of the year. Instead, borrowers play a version of Russian roulette, whereby the revolver is spun at least 12 times a year, or every time that a loan installment is due. We don't want to overstretch the Russian roulette analogy, but it really explains it quite well: with multiple elimination rounds in a year, what will be the odds of taking a bullet at each round, so that at year-end you would end up with 5 defaulters per 100 borrowers? In the default/no-default game, just like in Russian roulette, only survivors get to spin the revolver again. So, we can say that the 5% 1-year PD equates to a 95% cumulative survival rate after 12 monthly elimination rounds. With this, we can find the equivalent default rate at each monthly elimination such that:

$$(1-PD_{monthly})*(1-PD_{monthly})\dots *(1-PD_{monthly}) = (1-PD_{annual})$$

$$PD_{monthly}=1-(1-PD_{annual})^{(1/12)}$$

This is true because conditional probabilities are concatenated by multiplication: Only on the condition that you did not default in the first

> monthly round, do you get to spin the revolver again. And only the survivors of the second round get to play the third game etc.

If you are not familiar with the exponent notation in Excel, the **^** sign means "**to the power of**". And to the power of 1/12 is equivalent to taking the 12th root of something. So, have you typed it into Excel, yet? The monthly default probability that would be equivalent to a 5% annual default rate is: 0.427%.

Please keep Excel open, we want to look at the expected loss of a single loan under the relationship **EL=PD*EAD*LGD**. We will use the random number generator in Excel to do a simulation of the loan loss.

We copied in below the first line from the worksheet M4.1_Ex2_DefaultRate. The Loss is the result of multiplying the binary default variable with the EAD and LGD values. All three are set up as random variables.

Borrower No.	Loan Balance Outstanding	DefaultYes=1	EAD	LGD	Loss
1	1,000.00	1	1,139.52	0.485773	553.55

Figure 2: Screenshot of Datasheet in M4.1_Ex2_DefaultRate

Here is how we did this: The function **=RAND()** produces realizations of a random variable that are distributed with equal probability in the interval [0;1]. So, in order to get 5% defaults and 95% no defaults, we should write **=IF(RAND()>0.05,0,1).** All **=Rand()** values are refreshed with a new random result every time you save or recalculate the spreadsheet. You can trigger a refresh manually with the F9 button. So, please play with it a little and watch how the results change as you keep pushing the F9 button.

We also want to set up **EAD as a random variable**. Suppose we know that EAD should vary in the range of 110% to 130% of the amount currently outstanding. If the current balance is in cell B2, we can write **=(1.1+RAND()*0.2)*B2.**

For the **LGD random variable**, we got a bit more fancy: assume that we know from many years of lending experience to this type of market and with this type of collateral that LGD varies narrowly around the 60% average. This would be a case for using a normally distributed LGD variable with a mean of 60% and a narrow standard deviation of just 10% up and down from the 60% mean. We can generate random realizations of such a LGD variable by putting the equally distributed random numbers from **RAND()** through the inverse normal distribution function. This is done such that the **RAND()** value becomes the probability, for which we look up the limit along the x-axis that will give us that percentage value of cumulative probability in the normal distribution:

LGD =NORMINV(RAND(),0.6,0.1)

However, this formula could occasionally produce LGD values above 100% and below 0%, which make little sense. We should therefore cut off any excess over 100% or below 0% by using the **=MIN()** and **=MAX()** functions in Excel. That's why in the cell E2 we wrote:

LGD =MIN(MAX(NORMINV(RAND(),0.6,0.1),0),1)

In cell F2 you then find the **random loss amount** for the single loan as a function of the Default EAD and LGD random variables. If you push F9 100 times you will get a non-zero loss about 5 times.

Now copy and paste line 2 another 999 times below the first loan to get one thousand identical and independent individual loans that all follow the same random behavior of annual loss.

You can count the total number of defaults in the portfolio and obtain the random default rate for the portfolio in H2 as:

Portfolio Default Rate = SUM(C2:C1001)/1000

As you keep refreshing the random variables, you will see that the portfolio default rate varies around the 5% mark. Thus the PDs of the individual loans add up straight through to the portfolio level. The **portfolio probability of default is the arithmetic average of the individual PDs of each loan**.

Try it: Change the PD in the first 300 loans to 3%. We would expect a portfolio default rate of: (300*3%+700*5%)/1000 = 4.4%. And indeed, the Portfolio default rate now varies around the 4.4% mark as we push F9 repeatedly.

More generally speaking, the Excel simulation shows that the **expected losses from the individual loans simply add up to the expected value of the portfolio loss variable**.

However, for the other "moments" of the probability distribution, particularly the standard deviation of the portfolio loss distribution, we cannot just add the individual values to find the portfolio-level equivalent.

Referring back to our definitions of risk in Unit 1, you will remember that the **average expected portfolio loss is not "risk"**. This level of loss is pretty much a certainty and should be priced in and charged to borrowers by way of a mark-up on the interest rate. **The measure of risk is in the deviation from that expected average loss.** One should be concerned about the risk of having a bad year where portfolio losses are 10 times the expected value. This risk can be measured by the average deviation from the average portfolio loss, which is called the Standard Deviation.

> ⇨ *In statistics and probability theory,* ***standard deviation,*** *often represented by the symbol sigma σ, shows how much variation or "dispersion" exists from the average (mean, or expected value). The standard deviation of a random variable, statistical population, data set, or probability distribution is the square root of its variance.*

The standard deviation of a probability distribution can be estimated by the standard deviation S on a sample of N observations as:

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2}$$

The formula includes the common **sample bias correction** in that we take the average of the squared deviations from the mean by dividing by (N-1) and not by N. This is because we already calculated the mean on the same sample, hence the degrees of freedom are N-1. The sample standard deviation is available in Excel with the formula **<u>=STDEV()</u>**

Coming back to the **standard deviation of aggregate portfolio losses** as the measure of credit risk: Consider our loan portfolio in M4.1_Ex2. Instead of 1,000 loans with $1,000 each outstanding, imagine we just had a single loan to a single borrower with a balance of 1 million and all else equal: same PD, same EAD and LGD random variables. The expected loss is the same: 5%* 1.2 * 1 million * 0.6 = 36,000. Yet, the standard deviation around this expected portfolio loss will be smaller, if the portfolio is broken down into 1,000 individual loans rather than one big one. This is the famous subadditivity of risk and a result of diversification. This simply means that in a portfolio all borrowers will never default (or not default) at the same time, which prevents extreme deviations from the expected incidence of default. **We measure diversification in a portfolio by the degree of correlation**. Correlation is essentially a measure for the effect that if one defaults, how many of the other 999 in the portfolio will also default.

Under **perfect "positive" correlation**, all 1,000 loans will always default or not default together. This is identically risky as having lent the full one million to just one borrower; hence the standard deviation of the portfolio loss would be the sum of the standard deviations of the individual loan losses. In every other case, other than perfect positive correlation, the standard deviation of the portfolio loss will be smaller than the sum of the individual standard deviations. This is how diversification reduces risk.

Now, we would like to demonstrate the effect of portfolio diversification in an Excel simulation. We start with the same elementary loan as before:

- Balance outstanding: 1,000
- PD: 5% p.a.
- EAD: equally distributed in the interval [1.1; 1.3] * balance outstanding
- LGD: Normally distributed with mean 60% and standard deviation 10%, values below 0 and above 100% replaced by those limits

In order to keep it within the line limitation for Excel 2003 users, we copy that elementary loan with the underlying random variables 64,000 times into the same sheet. We will group these 64,000 loans into 128 parallel observations of the same portfolio consisting of 500 loans. Every 500 lines we sum the realized losses of the 500 preceding loans. That gives us 128 random snapshots of the aggregate portfolio loss with the benefit of diversification across 500 loans.

Loan Balance Outstanding	DefaultYes=1	EAD	LGD	Loss	Portfolio Losses	STDEV Portfolio Losses	STDEV Indiviual Loan Loss *500
1,000.00	0	1,253.37	0.688478	-		3,327.82	79,056.57
1,000.00	0	1,155.09	0.490701	-			
1,000.00	0	1,149.76	0.620816	-			
1,000.00	0	1,288.84	0.601142	-			

Figure 3: Screenshot from M4.1_Ex3_PortfolioLoss.xlsx

The standard deviation among these 128 observations of portfolio loss is in the $3,600 range, exactly 3,327.82 in this particular snapshot above. In parallel, we also calculated the standard deviation across all 64,000 individual loans along the column "Loss". The per loan standard deviation varies around the $160 mark. If you added these standard deviations up for all 500 loans in the portfolio, you would get a sum of standard deviations of around 80,000, i.e. 79,056.57 in this particular observation. The sum of standard deviations in I2 is more than 20 times larger than the actual standard deviation of the diversified portfolio loss in cell H2. This is subadditivity at work.

The way we set up the portfolios in PortfolioLossM4.1_Ex3 is actually a very specific situation among the possible diversification scenarios. The defaults on each of the 500 loans in the portfolio are **entirely independent from each other**, such that the outcome (default/no default) on one loan is not at all influenced by the fact that any of the other loans have defaulted or not. Perfect independence like this means entirely uncorrelated outcomes, i.e. a correlation coefficient of zero.

In our next exercise **M4.1_Ex4_PortfolioCorrelation,** we use the same set-up of 128 observations on a loan portfolio of 500 loans as in M4.1_Ex3. In order to make the file smaller and concentrate just on the impact of the default variable, we replace the EAD and LGD random variables by their fixed expected values. Now, instead of independence between the loan defaults, we assume perfect positive correlation, i.e. a correlation coefficient of 1.

Perfect positive correlation is easy to simulate. If one of the 500 loans in the portfolio defaults, they all default. So, if we know the outcome of the default variable for the first loan, we know the result for the next 499. Hence, we set the default variables in the 499 other loans equal to the variable in the first loan. This set-up we repeated 128 times, to simulate the outcome of 128 independent years for this perfectly correlated portfolio. And here, no surprise, the risk is fully additive: the standard deviation of the portfolio losses is equal to the sum of the standard deviations of the losses on the 500 individual loans. In the snapshot below, 82,174,55 is roughly equal to 81,853.57. If you keep pushing F9 on the sheet Positive Correlation in M4.1_Ex4, you will see that the two observed standard deviation values in the cells H1 and I1 always come out close to each other.

0	Loan Balance Outstanding	DefaultYes=1	EAD	LGD	Loss	Portfolio Losses	STDEV Portfolio Losses	STDEV Indiviual Loan Loss *500
1	1,000.00	1	1,200.00	0.6	720.00		82,174.55	81,853.57
2	1,000.00	1	1,200.00	0.6	720.00			
3	1,000.00	1	1,200.00	0.6	720.00			
4	1,000.00	1	1,200.00	0.6	720.00			
5	1,000.00	1	1,200.00	0.6	720.00			
6	1,000.00	1	1,200.00	0.6	720.00			
7	1,000.00	1	1,200.00	0.6	720.00			
8	1,000.00	1	1,200.00	0.6	720.00			

Figure 4: Screenshot from M4.1_Ex4_Portfolio Correlation.xlsx

In the second sheet in the same file for M4.1_Ex4, we now try to simulate a **strongly negatively correlated portfolio** of 500 loans, again with 128 parallel observations of the same portfolio. This is a little tricky to do in Excel, but we think we figured out a reasonable approximation. Each one of the 500 loans in the portfolio still has a standalone PD of 5% p.a., but this time the portfolio displays a high degree of compensating negative correlation. This would be the case, if for each loan that we make, we try to find another borrower who is in an opposing business, which does best when the first business defaults. Imagine always making one loan to a major corporation and another to a bankruptcy law firm headquartered in the same town. Both have a relatively low standalone PD and are definitely rather negatively correlated. If the corporation goes bankrupt, the bankruptcy lawyers are definitely not going to default. If the lawyers default, it is probably because corporate business in the town is doing very well and there have not been any major bankruptcies in the area in a long time.

In exercise M4.1_Ex4, **sheet Negative Correlation**, we tried to implement this idea as follows: We look at the portfolio as 250 pairs of opposing, negatively correlated loans. All have a standalone PD of 5% but loan 1 is matched with loan 251, loan 2 is matched with loan 252 etc. If borrower 251 defaults, borrower 1 certainly has a good year and does not default. This should also work in the other direction, but if you type the reverse relationship into the same formula between the two loans you naturally get a circular reference. So, for the reverse relationship we just shifted the pairs by one loan: the reverse effect is now between loan 1 and loan 252, loan 2 and loan 253 etc.

The default function in Cell C2 for Loan 1 is

=IF(C253=1,0,IF(RAND()>0.05,0,1))

and the reverse reference is found from cell C252 onwards for loan 251 as follows:

=IF(C2=1,0,IF(RAND()>0.05,0,1))

In plain English, the content of cell C252 just above says: if the paired loan number 1 defaults, this loan number 251 will certainly not default, otherwise loan 251 can still default with 5% PD. If we copy this reference pattern among the 250 loan pairs across each of the 128 portfolios in the worksheet, we can see that there is a **further reduction in the standard deviation of the portfolio losses** relative to the portfolio losses in the base case with independent defaults.

In the screenshot below from M4.1_Ex4 the standard deviation of the 128 observed portfolio losses is 2,987.95 versus 3,501.06 in the independence case and 76,418.92 under full positive correlation.

Loan Balance Outstanding	DefaultYes=1	EAD	LGD	Loss	Portfolio Losses	STDEV Portfolio Losses	STDEV Indiviual Loan Loss *500	Number of Defaults in 64,000 Loans
1,000.00	0	1,200.00	0.6	-		2,987.95	76,418.92	4.73%
1,000.00	0	1,200.00	0.6	-		Compared to Independence:		
1,000.00	0	1,200.00	0.6	-		3,501.06		
1,000.00	0	1,200.00	0.6	-				
1,000.00	0	1,200.00	0.6	-		Perfectly negative Correlation		
1,000.00	0	1,200.00	0.6	-		Correlation Coefficient:	-100%	
1,000.00	0	1,200.00	0.6	-				

Figure 5: Screenshot from M4.1_Ex4_PortfolioCorrelation, sheet Negative Correlation

Now, that we have a general idea about the "**EL=PD*EAD*LGD** way of thinking", we should take a more detailed look at the real world of microcredit and SME lending. PDs are not just simply "known" like the odds of rolling a six in a game of dice. Most of this Unit is about how, when and why a borrower default might occur, so that we can better estimate the probability of default at time of making the loan decision. We will also have another chapter about the finer details of LGD estimation and collateral, but understanding default is the big deal in credit risk, of course.

Chapter 2 - Review Questions

1. Define the expected loss in loan portfolio. What are its components?
2. All else equal, will increasing the effective rate charged on a microcredit product tend to increase or reduce LGD?
3. Name two reasons why EAD > balance outstanding is a plausible assumption in microcredit.
4. Default is generally modeled as what kind of a random variable: continuous, binary, or dummy variable?
5. How do you convert an annual PD into an equivalent calendar quarter PD? By dividing by four?

6. What does the subadditivity of risk refer to: the addition of expected losses in a portfolio or the aggregation of the standard deviation of the losses?