NEXIALOG
CONSULTING

UNIVERSITÉ PARIS 1
PANTHÉON SORBONNE

Risque de Crédit :
PD Forward Looking

Armand L'Huillier

Tel : 06.73.03.49.61
Mail: alhuillier@nexialog.com

THINK SMART ACT DIFFERENT

NEXIALOG
CONSULTING

Plan du cours du 4 décembre

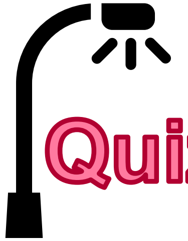
Cours PD FL

01 **Méthodologie de A à Z**

02 **Appli PDFL**

Challenge Talk

03 **Sujet, questions, conseils**



Quizz?

Tel : 06.73.03.49.61
Mail: alhuillier@nexialog.com

Cours PD Forward Looking

Cours PD FL

Tel : 06.73.03.49.61
Mail: alhuillier@nexialog.com

THINK SMART ACT DIFFERENT

Cours PDFL

Méthodologie de A à Z

1 - Les notions fondamentales

Le « Forward looking »

Définition

- Un test de résistance bancaire, ou « stress test », est un exercice consistant à **simuler des conditions économiques** et financières extrêmes mais plausibles afin d'en étudier les conséquences sur les banques et de **mesurer leur capacité de résistance** à de telles situations. Mots clés : Capital économique.

Contexte

- Les stress tests ont été mis en place par les banques centrales et les autorités en charge de la supervision bancaire à la fin des années 1990. La crise asiatique de 1997 avait mis en évidence l'importance de la détérioration des facteurs économiques dans le déclenchement des crises. **Tous les secteurs font des stress tests** (assurance, énergie…).

Réglementation

vs l'estimation des paramètres bâlois.

- But : Aller plus loin que la Réglementation IAS 39 (où seuls les événements passés et les conditions actuelles pouvaient être considérés pour mesurer les pertes de crédit → comptabilisation retardée des baisses de crédit)
- En plus de prendre en compte les événements passés et les conditions actuelles, la nouvelle norme **IFRS 9 (entrée en vigueur en 2018)** exige que les informations prospectives soient utilisées pour mesurer les **pertes de crédit attendues (ECL)**(Pour les ECL lifetime).
- Autres besoins : Pilier II, ICAAP, simulations, stress…

Objectifs

- Les exercices de stress test visent donc à mesurer l'impact du choc macro-économique sur le risque de crédit porté par les banques, la valeur de leurs actifs et in fine sur leur **ratio de solvabilité** et compte de résultat.
- Ils ont pour objectif d'évaluer la capacité des banques à faire face à des crises économiques éventuelles ou à des contextes économiques dégradés sur un horizon de temps donné (horizon de 3 à 5 ans).
- Des actions correctives sont prises lorsque les résultats des stress tests montrent une incapacité des banques à faire face à de potentielles récessions économiques.

Econométrie : l'économie a un impact sur les défauts ?

Quelles données pour étudier la macroéconomie ?

- 3 scenarios
- SOURCE : Macro-financial scenario for the 2023 EU-wide banking sector stress test (European Systemic Risk Board (ESRB).
- Souvent, les banques font un modèle par pays : la variable du taux de chômage peut être significative en France, mais pas en Espagne. Et le coefficient peut aussi être different selon les pays.

Séries macroéconomiques

Base issue de la BCE

Série du taux de défaut DRt

Modélisée plus tôt par l'équipe de risque de crédit de la banque

Calculs

Variables explicatives :
- Séries macro Lagged
- Taux de défaut Lagged

Estimation d'une régression DRt = Xβ + ε

Calcul des paramètres Zt et ρ ……

Avec Merton-Vasicek

Matrice de migration Through the Cycle (TTC)

Matrice de référence : Une par année

Calculs

Historique détaillé des défauts & Notes

Donné par la banque

Matrice de migration Point in Time (PIT) → Matrices projetées

Stress sur les matrices TTC grâce à la régression
→ Résultant en une matrice par scénario

Calculs des PD Forward looking

NEXIALOG CONSULTING

THINK SMART ACT DIFFERENT

Cours PDFL

Méthodologie de A à Z

2 - Projection du taux de défaut

Les prétraitements

Zoom sur la stationnarité

Un processus Xt est stationnaire au second ordre si : $E(X_t) = m$ (Ne dépend pas du temps); $V(X_t)$; $Cov(X_t; X_{t+h}) = \Gamma(h)$ Avec $\Gamma(h)$ la fonction d'auto - covariance du processus.

Traitement du trend et/ou de la partie saisonnière : souvent présentes. Plusieurs manières d'observer la non-stationnarité : Graphique de la série, Corrélogrammes, test portemanteau, Test **ADF** (aug. Dickey Fuller, test sur les racines unitaires), **Philippe Perron, KPSS.**

Visualisation, tests, transformations

- Lissage
- Saisonnalité

Des variables macros bien choisies

- On prend les **LAG (retards)** : le modèle aura toujours un temps de retard (t-1)
- Beaucoup de variables candidates : lag **mensuels**, **annuels**, allant de t-1 à t-X …
Elles doivent être **stationnaires**: (comme dans toutes les régressions) elles le sont rarement → **traitement** pour les garder
- On se cantonne aux variables des économistes de la banque : pas sur internet normalement.

Construction des variables

Les variables sont des taux, des séries brutes… on peut souhaiter les transformer en taux de **variation**.

- Trimestrielles
- Annuelles
- …

Corrélations

Pearson peut suffire en général. Mais en banque **Kendall est préféré :** il est stable sur un échantillon de petite taille.
Il fonctionne par paires. Corrélation si $\tau \in [0.5,\ 1]$ en valeur absolue.
Utile pour retirer les variables trop corrélées entre elles.

La série du taux de défaut n'est pas stationnaire

Or cette série est la variable endogène à modéliser. On NE peut PAS la différencier pour faire disparaître la **non-stationnarité**, ceci nous empêcherait de modéliser le Zt et le paramètre ρ.

Donc on utilise une astuce statistique…

→ RDV slide suivante les apprentis data scientists ;)

L'astuce en dernier recours pour régler le problème de non-stationnarité

D'où vient le problème ?

- La série du taux de défaut (DR) n'est généralement pas stationnaire par construction. En effet, deux observations consécutives de la série DR partagent trois DR trimestriels identiques, ce qui entraîne un problème de forte autocorrélation.

Solutions en Série Temp

On a donc recours à différentes méthodes voici ci-dessous une liste non exhaustives :

- Méthode de cointégration (deux séries non stationnaire → une stationnaire);
- Modèle à changement de régime;
- Détection des ruptures dans la série (Avoir assez d'historique) ;
- Transformation en différence partielle première de DR.

Résolution

$$DR_t - \beta \times DR_{t-1}$$

Un processus de différenciation est appliqué à la série DR afin d'atteindre la stationnarité et d'éviter ainsi une régression MCO erronée. Cette nouvelle série peut être interprétée comme une **différenciation partielle** de la série DR. Le processus de sélection du β optimal consiste à **tester différentes valeurs** en gardant comme objectif d'avoir une valeur bêta qui:

- conduit à une série stationnaire,
- est faible (ce qui conduirait à une contribution plus élevée des macro-variables)
- a une capacité d'ajustement (R2) qui soit élevée
- a un bon ajustement hors échantillon sur les crises économiques passées
- est homoscédastique et sans autocorrélation sérielle.

Un algorithme a été créé pour sélectionner l'intervalle optimal grâce aux tests PP et KPSS. But : garder la plus faible valeur de bêta tout en conduisant à la zone de rejet la plus probable de ces tests. Cet algorithme repose sur la procédure de Newey & West (1987).

$$DR_t = \frac{Number\ of\ counterparties\ falling\ into\ default\ between\ t-4\ and\ t}{Number\ of\ performing\ counterparties\ at\ t-4}$$

Comment modéliser la série DR ?

- Une fois qu'on a les variables : DRt-1, DRt-2, DRt-3…. Les variables macroéconomiques lagged : taux de chômage t-1, taux de chômage année-1, PIB…. Et qu'on a vérifié toutes les hypothèses : peu de corrélation entre les variables explicatives, l'intervalle de Beta qui satisfait la stationnarité de DRt… Alors on peut modéliser la série DRt.

ADL
regression

Le modèle traditionnel pour modéliser une série temporelle dans ce cas : ADL (Autoregressive Distributed Lag)

S'estime par la méthode des MCO.
Modèle simple, basique, interprétable. Mais cette structure souffre de multicolinéarité récurrente.

« ElasticNet is for the best »

$$DR_t = \alpha + \sum_{s=0}^{p} \emptyset_s DR_{t-s} + \sum_{s=0}^{q} \theta_s x_{t-s} + \varepsilon_t$$

Deux méthodes d'estimation sont donc possibles :
- Incorporer une variable macro-économique par tâtonnement et estimer les modèles un à un.
  → méthode peu efficace, non robuste, non optimale.
- Utiliser les méthodes ensemblistes

BMA
model

Bayesian model average : Méthode de sélection de modèles proposé par la BCE dans la STAMPE.

- La méthode ensembliste choisie pour remédier au problème de multi-collinéarité est le BMA. Il réduit l'incertitude liée à la présence de variables.
- Zoom dans la slide suivante…

Diversity is a luxury

Autres modèles

- Dans l'idée de proposer des modèles challengers, on peut utiliser des modèles moins interprétables : Random Forest regression…
- Permet de comparer les perfs.

Estimer le modèle Bayésien

- L'estimation des $2^{\text{nombre\_de\_variables\_explicatives}}$ sous-modèles est la première étape pour arriver au modèle final.

- Nous avons une série de modèle : M = $\{Mi,..,Mk\}$
- Celle-ci va être donnée en fonction des différentes combinaisons possibles : k = 2^{P} → itertools sur python.
- Le but est de calculer la probabilité a posteriori de chaque sous-modèles, de les sélectionner et ensuite faire leur moyenne pondérée.
- On note la probabilité a posteriori : $P_r\,(M_k\,|Données)$.
- $\mathrm{P_r(M_k|Données)} = \frac{\mathrm{P_r(Données|M_k)\ P_r(M_k)}}{\sum_{l=1}^{k} \mathrm{P_r(Données|M_l)\ P_r(M_l)}}$ représente la probabilité a posteriori.
- $\mathrm{Pr(Données|Mk)} = \int \mathrm{Pr(Données|\theta k, Mk)\ Pr(\theta k|Mk) d\theta k}$ représente la vraisemblance marginale du modèle $\mathrm{M_k}$ et θ_k le vecteur de paramètres du modèle Mk.
- $\mathrm{Pr(\theta k|Mk)}$ la densité a priori de θ_k sous le modèle $\mathrm{M_k}$ et $\mathrm{Pr(Données|\theta k, Mk)}$ la vraisemblance.
- La probabilité marginale de chaque modèle permet de repondérer les probabilités a priori $\mathrm{P_r(M_k)}$.
- Nous pouvons utiliser ces probabilités comme poids pour faire des inférences et obtenir les moyennes pondérées des quantités d'intérêt.
- On note Δ la quantité d'intérêt. Donc la probabilité a posteriori de Δ sachant que j'ai observé mes données résulte de :

$$P_r(\Delta\,|Données) = \sum_{l=1}^{k} P_r(\Delta\,|M_k, Données)\ P_r(M_k|Données)$$

$$E(\Delta\,|Données) = \sum_{k=1}^{K} \hat{\Delta}_k\ P_r(M_k|Données)$$

$$V(\Delta\,|Données) = \textstyle\sum_{k=1}^{K}\{V[\Delta|M_k, Données] + \hat{\Delta}_k^2\}P_r(M_k|Données) - E[\Delta|Données]^2$$

On sélectionne les sous-modèles les plus pertinents, on fait la moyenne pondérée de coefficients. On sélectionne les sous-modèles en fonction des probabilités a posteriori (méthode complexe : il faudrait calculer C avec la fenêtre d'Occam, Adrien Raftery). La proba a posteriori se calcule à partir du BIC : $p_r(M_k|D) = \frac{e^{-0.5*BIC_k}}{\sum e^{-0.5*BIC_k}}$
Ce qui va donner la probabilité a posteriori comprise entre 0 et 1 de chaque modèle M_k. Deux choix peuvent s'offrir à nous : garder le modèle avec les poids associés ou <u>sélectionner les variables avec la plus forte probabilité a posteriori</u>.

Hypothèses : $E(\varepsilon) = 0$, pas d'autocorrélation des ε, homoscédasticité, pas de multi−colinéarité

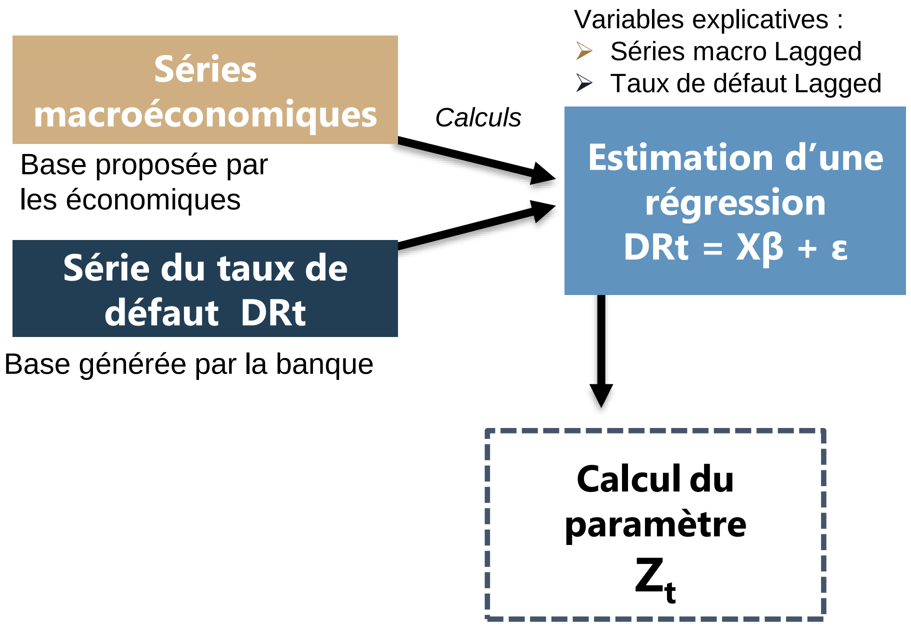
Métriques : U de Theil, Intervalle de confiance, RMSE, MSE, MAE, R2.

Besoin métier : Significativité des coefficients, sens économique !

Finalité du modèle DRt

- Le modèle permet de projeter le DRt à horizon 1 an, 2 ans…

- On prend les scénarios BCE : Adverse, Central et Favorable.
- Chaque année les économistes de la BCE publient des prévisions économiques. Celles-ci sont utilisées dans les modèles pour anticiper une hausse ou une baisse du taux de défaut selon les prévisions.
- En plus, les banques ont des équipes d'économistes qui fournissent plus de données (plus de variables, des variables sur des secteurs spécifiques, sur des zones géographiques particulières).



THINK SMART ACT DIFFERENT

Cours PDFL

Méthodologie de A à Z

Matrices de migrations

Les matrices de migration

A quoi ressemble une matrice de migration ?

- Description empirique des migration d'une note à une autre.
- *Conditionnelle à une classe de notation donnée en date T, la matrice de migration décrit les probabilités (pour le débiteur) d'être dans l'une des différentes classes de notation en date T+1. Elle décrit la distribution des probabilités de notation à la date T+1 étant donné la notation de départ en date T.*
- Les matrices de migration annuelles synthétisent l'évolution des notations des contreparties saines sur **l'horizon annuel**. Voici donc le calcul du taux de migration de rating à t+12 mois:

$$\boldsymbol{Taux\ de\ migration\ rating_{i,j,t}} = \frac{\boldsymbol{Nombre\ de\ contreparties_{i,j}}}{\boldsymbol{Nombre\ total\ de\ contreparties_{i,t}}}$$

- Avec,

$Taux\ de\ migration\ rating_{i,j,t}$: Taux de passage du rating i vers le rating j à la date t

$Nombre\ de\ contreparties_{i,j}$: Volume de contreparties qui migrent du rating i vers le rating j

$Nombre\ total\ de\ contreparties_{i,t}$: Volume total de contreparties ayant un rating i

Probabilité de passer de la note 1 à 6.

	1	2	3	4	5	6	7	8	D
1	90,397%	8,532%	0,868%	0,167%	0,033%	0,001%	0,001%	0,000%	0,000%
2	1,265%	89,829%	7,808%	0,791%	0,182%	0,041%	0,006%	0,005%	0,072%
3	0,082%	3,056%	89,932%	5,915%	0,744%	0,128%	0,030%	0,009%	0,102%
4	0,044%	0,305%	4,748%	88,255%	5,342%	0,846%	0,141%	0,016%	0,303%
5	0,008%	0,093%	0,527%	6,684%	82,843%	7,681%	0,657%	0,076%	1,433%
6	0,007%	0,054%	0,170%	0,660%	6,558%	81,429%	6,337%	0,599%	4,187%
7	0,000%	0,023%	0,031%	0,201%	0,871%	9,529%	71,462%	4,291%	13,592%
8	0,000%	0,029%	0,126%	0,068%	0,577%	3,678%	9,212%	58,606%	27,704%
D	0,000%	0,000%	0,000%	0,000%	0,000%	0,000%	0,000%	0,000%	100,000%

La matrice TTC, la référence

Construire la matrice de migration TTC (through the cycle), la matrice

- La construction de la matrice de référence, matrice TTC (Through-The-Cycle), est identique à celles des matrices mensuelles/trimestrielles. Elle est construite en tenant compte des historiques de rating de toutes ces dernières.
- Le calcul du taux de migration de rating à t+12 mois se fait selon la formule suivante:

$$\boldsymbol{Taux\ de\ migration\ rating_{i,j}} = \frac{\Sigma_{i=date\ de\ dbt\ d'historique}^{date\ de\ fin\ d'historique} \boldsymbol{Taux\ de\ migration\ rating_{i,j,t}}}{\boldsymbol{N}}$$

- Avec,
- $Taux\ de\ migration\ rating_{i,j}$: Taux de passage du rating i vers le rating j à la date t
- N: Nombre total de matrices

t + 12 mois

t

Rating	3	4	5	6	7	8	9	10	D
3	90%	10%	3%	…	…	…	…	…	0,3%
4	18%	56%	15%	9%	…	…	…	…	0,5%
5	9%	…	45%	4%	1%	…	…	…	0,6%
6	4%	…	…	38%	2%	…	…	…	1,7%
7	0,3%	…	…	…	30%	8%	…	…	2,3%
8	…	…	…	…	…	28%	…	…	4,5%
9	…	…	…	…	…	…	17%		11%
10	…	…	…	…	…	…	…	10%	40%

Taux de défaut, DR_t^{obs}

La matrice TTC doit respecter des propriétés:

- Toute probabilité de migration est strictement positive.

$$\forall\ i,j;\ \ p_{i,j} > 0$$

- La probabilité de passer en défaut est plus élevée pour les mauvais ratings que pour les bons ratings

$$j < i;\ \ p_{i,Def} > p_{j,Def}$$

- La probabilité de passer à un rating proche est plus élevée que la probabilité de passer à un rating éloigné

$$p_{i,i+1} > p_{i,i+2} > p_{i,i+3} > \ \dots .p_{i,i+n}$$
$$p_{i,i-1} > p_{i,i-2} > p_{i,i-3} > \ \dots .p_{i,1}$$

La MM TCC Cumulée

Pour projeter le taux de défaut avec les paramètres de Merton-Vasicek, on aura besoin des matrices cumulées.

	1	2	3	4	5	6	7	8	D
1	90,397%	8,532%	0,868%	0,167%	0,033%	0,001%	0,001%	0,000%	0,000%
2	1,265%	89,829%	7,808%	0,791%	0,182%	0,041%	0,006%	0,005%	0,072%
3	0,082%	3,056%	89,932%	5,915%	0,744%	0,128%	0,030%	0,009%	0,102%
4	0,044%	0,305%	4,748%	88,255%	5,342%	0,846%	0,141%	0,016%	0,303%
5	0,008%	0,093%	0,527%	6,684%	82,843%	7,681%	0,657%	0,076%	1,433%
6	0,007%	0,054%	0,170%	0,660%	6,558%	81,429%	6,337%	0,599%	4,187%
7	0,000%	0,023%	0,031%	0,201%	0,871%	9,529%	71,462%	4,291%	13,592%
8	0,000%	0,029%	0,126%	0,068%	0,577%	3,678%	9,212%	58,606%	27,704%
D	0,000%	0,000%	0,000%	0,000%	0,000%	0,000%	0,000%	0,000%	100,000%

Σ

Somme des probabilités de toutes les migrations de la classe 1 vers les classes strictement supérieures à 1 i.e.

$$\sum_{i=2}^{D} \boldsymbol{ProbaMigrationMarg_i^1} = \mathbf{9{,}602\%}$$

probabilité de migration cumulées de la classe 1 vers les classes strictement supérieures à la classe 1

2. Matrice TTC cumulée

	1	2	3	4	5	6	7	8	D
1	100,000%	9,602%	1,070%	0,202%	0,035%	0,002%	0,001%	0,000%	0,000%
2	100,000%	98,734%	8,905%	1,097%	0,306%	0,124%	0,083%	0,077%	0,072%
3	100,000%	99,916%	96,860%	6,928%	1,013%	0,269%	0,141%	0,111%	0,102%
4	100,000%	99,956%	99,651%	94,903%	6,648%	1,306%	0,460%	0,319%	0,303%
5	100,000%	99,994%	99,901%	99,374%	92,690%	9,847%	2,166%	1,509%	1,433%
6	100,000%	99,994%	99,940%	99,770%	99,110%	92,552%	11,123%	4,786%	4,187%
7	100,000%	100,000%	99,977%	99,946%	99,745%	98,874%	89,345%	17,883%	13,592%
8	100,000%	100,000%	99,971%	99,845%	99,777%	99,200%	95,522%	86,310%	27,704%
D	100,000%	100,000%	100,000%	100,000%	100,000%	100,000%	100,000%	100,000%	100,000%

Probabilité de défaut* mesure TTC v/s mesure PiT

TCC

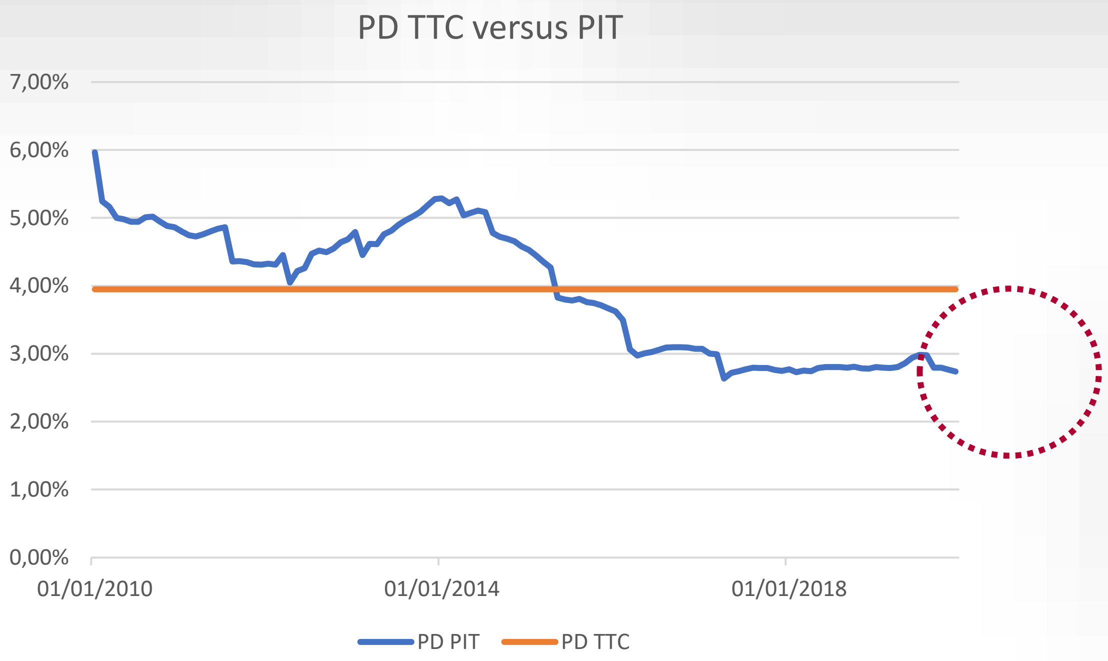
Mesure « Trough the Cycle » (TTC) : reflète la tendance long terme, indépendantes des conditions économiques courantes

- **Probabilités non conditionnelles à l'état de la conjoncture**
- **Calcul des exigences en fonds propres, pour limiter la procyclicité**

Point In Time (PIT)

La mesure Point-in-Time (PIT), reflète l'information disponible à une date donnée, reflète les conditions économiques courantes

- **Probabilités conditionnelles à l'état de la conjoncture**
- **Calcul des provisions IFRS 9**



Un objectif de la modélisation PD FL est de construire des matrices de migration PIT, en fonction des conditions économiques attendues.

* PD pour une note donnée. Le raisonnement s'applique pour toutes les notes.

NEXIALOG
CONSULTING

THINK SMART ACT DIFFERENT

Cours PDFL

Méthodologie de A à Z

Modèle de Merton-Vasicek

Modèle de Merton-Vasicek

Raisonnement

L'hypothèse principale du modèle est que l'évènement de faire défaut pour un client i dépend d'une variable Z qui suit une loi normale (0, 1) qui représente le risque **systémique**. Ksi représente le **facteur individuel** (suit aussi une normale (0, 1)).

$$Y_i = \sqrt{\rho} Z + \sqrt{1-\rho}\,\xi_i$$

Le paramètre Ro est la corrélation entre les différents facteurs de défaut des différents clients. On l'interprétera comme la sensibilité à l'environnement économique, donc la sensibilité du client à Z.
En découle la formule de Vasicek déjà utilisée dans la réglementation Bale II. La formule de la probabilité de faire défaut conditionnelle au paramètre Z :

$$PD_i(Z) = \Pr\left[\sqrt{\rho} Z + \sqrt{1-\rho}\,\xi_i Y_i \le \Phi^{-1}(PD_i) | Z\right] =$$

$$= \Pr\left[\xi_i \le \frac{\Phi^{-1}(PD_i) - \sqrt{\rho} Z}{\sqrt{1-\rho}}\right] = \Phi\left(\frac{\Phi^{-1}(PD_i) - \sqrt{\rho} Z}{\sqrt{1-\rho}}\right).$$

La probabilité conditionnelle de faire défaut peut être utilisée comme un proxy du taux de défaut futur. Z étant inconnu, elle est dépendante d'un facteur systématique à déterminer. Sur des portefeuilles très larges, on fait l'hypothèse que Ro est égal pour tous les clients.
Une fois que la distribution de Z sera estimée, il ne restera plus qu'à stresser le paramètre systémique Z et on obtiendra le taux de défaut futur.

Modèle de Merton-Vasicek

Paramètre 1 : ρ

Ce paramètre représente la sensibilité à la conjoncture économique.

La paramètre ρ est donné par la formule suivante :

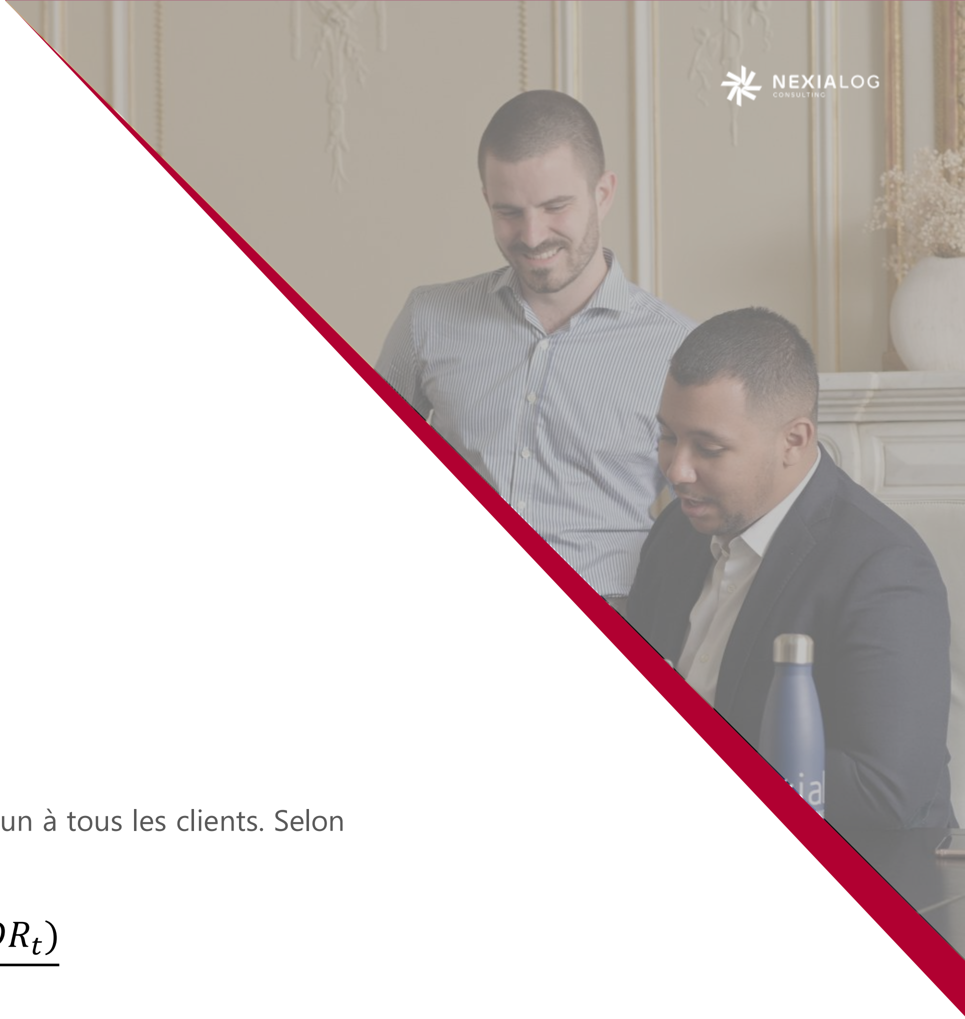
$$\rho = \frac{\mathbb{V}[\Phi^{-1}(DR^{historical})]}{1 + \mathbb{V}[\Phi^{-1}(DR^{historical})]}$$

Paramètre 2 : Z_t

Vasicek permet de construire Z_t. Ce paramètre représente la conjoncture macro-économique. Il est commun à tous les clients. Selon le modèle Vasicek, la série Z_t se construit avec la formule :

$$z_t = \frac{\Phi^{-1}(\overline{DR}^{historical}) - \sqrt{1-\rho} \times \Phi^{-1}(DR_t)}{\sqrt{\rho}}$$

- $\overline{DR}^{historical}$: la moyenne des taux de défaut historiques
- DR_t : le taux de défaut observé à l'instant t



NEXIALOG CONSULTING

THINK SMART ACT DIFFERENT

Cours PDFL

Méthodologie de A à Z

Stress sur les matrices de migrations

Matrices stressées, horizons 1 à 3 ans

Une matrice de migration de rating à 1 an par type de scénario macroéconomique (central, adverse, favorable) et par année de projection (horizon de 3 ans) sera construite. Cet horizon Forward-Looking de 3 ans correspond à l'horizon de projection maximal des variables macroéconomiques par les économistes de la banque. Au total, 3x3 matrices de migration de rating Forward-Looking PIT seront construites à cette étape.

Chaque matrice prend en compte l'impact des anticipations économiques sur la probabilité de migration d'une classe de rating à une autre. Ainsi, les probabilités de migration de rating Forward-Looking contenues dans cette matrice s'obtiennent via :

- les projections du paramètre Z ;
- les probabilités de migration de rating de la matrice de référence.

$$\forall \ j \geq i, \qquad p_{i,j}^{cum}(z_{proj}) = \Phi\left(\frac{1}{\sqrt{1-\rho}}\left[\Phi^{-1}(P_{i,j}) - \sqrt{\rho}\, z_{proj}\right]\right)$$

Avec :

- $z_{proj} = \frac{\Phi^{-1}(\overline{DR}^{historical}) - \sqrt{1-\rho} \times \Phi^{-1}(DR_{proj})}{\sqrt{\rho}}$
- $\overline{DR}^{historical}$: la moyenne des taux de défaut historiques
- DR_{proj} : le taux de défaut projeté à partir de modèle économétrique
- $P_{i,j}$: la probabilité de migration du rating i vers le rating j

Matrices cumulées

Input (pour chaque scénario) :

- les 3 matrices de migration de rating Forward-Looking PIT ;
- la matrice de migration TTC (référence).

La structure par terme de PD est construite grâce à la multiplication des matrices de migration de rating à 1 an sous l'hypothèse markovienne.

- Pour les 3 premières années, correspondant à la période couverte par les projections macroéconomiques, les matrices de migration de rating Forward-Looking PIT (M_{FWD_y}) , avec $y \in [1; 2; 3]$, sont utilisées.

Exemple avec la matrice de l'<u>année 2</u>:

La matrice de migration de rating cumulée à 2 ans est définie comme suit :

$$M_2^{cumul} = M_{FWD_1} \; M_{FWD_2}$$

$CEDF_2(rating = i)$ est la probabilité de défaut cumulée à 2 ans PIT pour le rating i :

$$CEDF_2(rating = i) = M_2^{cumul}[i, default]$$

- **Après l'horizon Forward-Looking ($y > 3$), la matrice de migration de rating TTC (M_{TTC}) est utilisée.**

Le calcul de la structure par terme repose sur un processus itératif :

Hypothèse: les transitions de rating respectent les propriétés des chaînes de Markov.

Implication: Multiplication des M_{FWD_i}

THINK SMART ACT DIFFERENT

Cours PDFL

Application

THINK SMART ACT DIFFERENT

Challenge Nexialog X Mosef

Sujet, questions, conseils

Challenge : Kick-off - 04 décembre

Partie 1 : Modélisation

Suivre les slides de présentation du projet d'Aryan du lundi 23 octobre.

Sujet : **Modélisation de la PB bâloise**

First steps –

Step 1 = Prendre connaissance des données : **identifier les natures des variables** (**1**.Borrowers' loan information, such as loan type and amount **2**.Borrower's credit history information, such as the number of financial accounts and delinquencies **3**.borrower's soft information), **data Viz** (sur la cible et les variables exogènes), faire **des tests de corrélation, d'indépendance**

Step 2 = faire un **choix** sur les variables intéressantes, appliquer un preprocessing adapté à la problématique.

Step 3 …

…

Partie 2 : Création d'une application

L'application (dash > streamlit) servira deux objectifs.

Présenter la méthodologie/le modèle. Application de votre modèle sur un client fictif; Prédiction de son défaut, sa classe…

First steps –

Faire le modèle de la partie 1 !

NB :

```
raw_data = pd.read_csv(path_file)
```

- Envoie de la base par mail.
- Les slides PDFL
- **Cours de suivis**