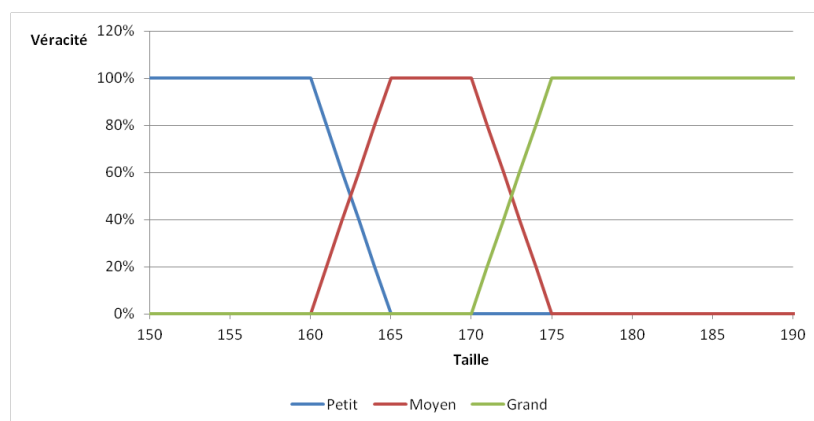


1. PRINCIPES

La logique floue est un type de modélisation qui s'intéresse à la prédiction d'une variable catégorielle Y « subjective » au sens où elle n'est pas objectivable : elle dépend de l'observateur (l'individu est « grand », « moyen » ou « petit »). Ce cadre sort de la statistique classique dans lequel la valeur de la variable Y est objectivable (« l'individu mesure 176 cm »). L'application de la logique floue revient à tenter d'appliquer un raisonnement proche de la pensée humaine :

- Les variables prédictives (comme la variable à prédire) sont catégorielles avec des modalités subjectives (« grand », « petit ») et non pas de données objectivables (176 cm). Ces variables catégorielles sont appelées « variables linguistiques ». Dans le cadre statistique usuel, la variable continue initiale (ici la taille en cm) peut être discrétisée pour donner des intervalles distincts, par exemple : « petit < 170cm < moyen < 180cm < grand ». La logique floue vise à prendre en compte les incertitudes qui existent au voisinage des seuils (due en partie à des principes de subjectivité).
- Une donnée peut appartenir à plusieurs modalités d'une même variable (un individu de 165 cm peut être considéré comme petit mais aussi comme moyen). Les classes définies ne partitionnent donc pas l'ensemble des possibles car elles peuvent se recouper.
- La logique floue intègre un ensemble de règles permettant d'attribuer (d'une manière logique) une sortie à une entrée.

La logique floue permet donc d'intégrer des systèmes experts dans des processus automatisés. Ce point constitue à la fois une force et une faiblesse de la logique floue. Le graphique de véracité suivant montre qu'un individu de 162 cm peut être considéré en logique floue comme étant petit à 60% et moyen à 40%



Au-delà de cette différence de principe, elle intègre également une prise en compte des interactions différentes de celle du monde probabiliste en redéfinissant les opérateurs logiques.

	Opérateurs flous de Zadeh	Opérateurs probabilistes
A ET B	$\min(\mu_A, \mu_B)$	$\mu_A \times \mu_B$
A OU B	$\max(\mu_A, \mu_B)$	$\mu_A + \mu_B - \mu_A \times \mu_B$
NON A	$1 - \mu_A$	$1 - \mu_A$

Tableau 1 : traduction mathématique sous la logique Floue

Dans le Tableau 1 : traduction mathématique sous la logique Floue, μ désigne la fonction de véracité pour les opérateurs flous qui est l'analogue de la mesure de probabilité en théorie des probabilités classiques.

On remarque que :

$$(A \text{ ET } B)_{\text{Probabiliste}} \leq (A \text{ ET } B)_{\text{flou}} \leq (A \text{ OU } B)_{\text{flou}} \leq (A \text{ OU } B)_{\text{Probabiliste}}$$

Développée à partir de 1965 par le professeur Lofti Zadeh de l'université de Berkeley dans un article fondateur qui en définit les principes (ZADEH, 1965), elle constitue une généralisation des ensembles classiques. Elle commence à être utilisée dans l'industrie, la médecine, la mise en place de système experts dans le milieu des années 70 puis verra son utilisation généralisée dans les années 90 (autofocus, autocuiseurs, systèmes autonomes mobiles, systèmes de décision, de diagnostic, de reconnaissance).

Son fonctionnement peut se résumer en trois grandes étapes :

- **La fuzzification** : transformation des variables en variables floues (aussi appelées variables linguistiques) en leur associant des lois de véracité (la variable taille est divisée en modalités « un individu de taille 162 cm est « petit » à 60%, « moyen » à 40% et « grand » à 0% »). Ce procédé s'apparente à la définition de lois *a priori* en statistiques bayésiennes, avec dans cet exemple une loi *a priori* (0,6 ; 0,4 ; 0). La différence dans ce cadre est que la somme des véracités n'est pas tenue de valoir 1.
- **L'inférence floue** : construction de règles (et de résultats) basées sur les variables linguistiques, attribution d'une véracité à chaque règle, puis agrégation des règles pour obtenir un résultat (linguistique) unique
- **La defuzzification** : passage d'un résultat linguistique à un résultat chiffré.

Ces différentes étapes sont reprises en détail ci-après.

2. LES TROIS ETAPES DE CONSTRUCTION D'UN MODELE DE LOGIQUE FLOUE

1. La fuzzification

Cette première étape consiste à transformer les variables (d'entrée et de sortie) en variables linguistiques :

- Pour chaque variable, on définit dans un premier temps l'univers du discours (i.e. la plage de valeurs que peut prendre la variable).
- La variable est ensuite découpée en catégories appelées variables linguistiques
- Une fonction (allant de 0% à 100%) permettant de définir pour chaque variable son pourcentage de véracité à l'affirmation : « l'observation est dans telle catégorie » est affectée à chaque catégorie

Cette étape est principalement réalisée sur la base d'observations statistiques (ou par apprentissage, supervisé ou non, pour regrouper les valeurs d'une variable en catégories homogènes) ou à dire d'expert. Les graphiques ci-après sont tracés en prenant l'exemple de la taille.

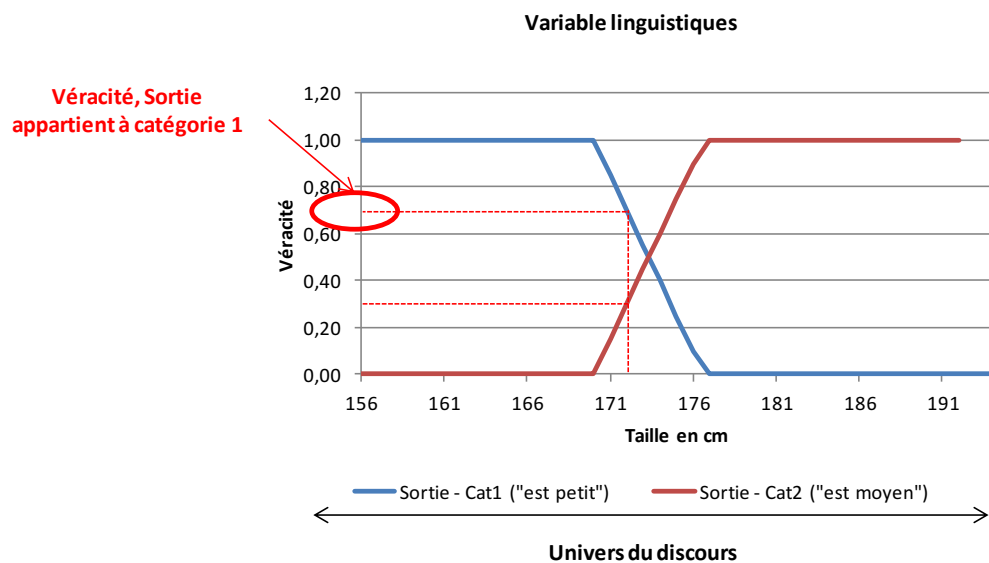


Figure 1 : Fonction de véracité

Ainsi, dire que la sortie est dans la catégorie 1 si la variable en sortie vaut 75% a une véracité de 70% et une véracité de 30% pour la catégorie 2.

- **Univers du discours** : taille en cm comprise entre 156cm et 194cm
- **Variable linguistique** : nom de la variable de sortie (par exemple la taille)
- **Valeurs linguistiques** : « Cat1 » (petit), « Cat2 » (moyen).

2. L'inférence floue

a. Construction d'un ensemble de règle

Sur la base des catégories précédemment réalisées, un ensemble de règles sont construites. Par exemple : « Variable 1 Catégorie 1 et Variable 2 Catégorie 1 ».

Une véracité pour chacune des règles est alors calculée. La construction de ces règles, principalement basée sur des « ET », des « OU » et des « SANS », se traduit mathématiquement de la sorte.

b. Matrice de décision

Chaque règle se voit attribuer une réponse au travers d'une matrice de décision.

Sortie	Var 2 Cat 1	Var 2 Cat 2
Var 1 Cat1	Cat1	Cat 2
Var 1 Cat2	Cat 2	Cat 1

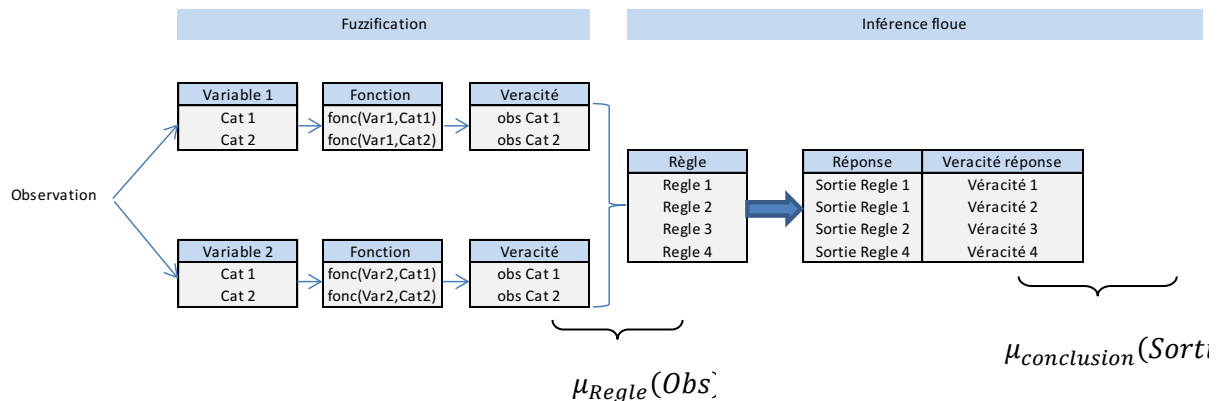
Tableau 2 : matrice de décision

Ceci pouvant se réécrire :

Règle (i)	Descriptif (R_i)	Règle de sortie ($Conclusion_{R_i}$)
1	Var 1 Cat 1 et Var 2 Cat 1	Sortie Cat1
2	Var 1 Cat 1 et Var 2 Cat 2	Sortie Cat2
3	Var 1 Cat 2 et Var 2 Cat 1	Sortie Cat2
4	Var 1 Cat 2 et Var 2 Cat 2	Sortie Cat1

Tableau 3 : matrice de décision, vue éclatée

A cette étape, une observation suit donc le parcours suivant :



c. Implication : calcul de la règle d'activation

Il reste à définir une règle d'activation afin d'obtenir une réponse unique. Cette étape s'appelle l'implication. Elle peut être effectuée au travers de deux règles :

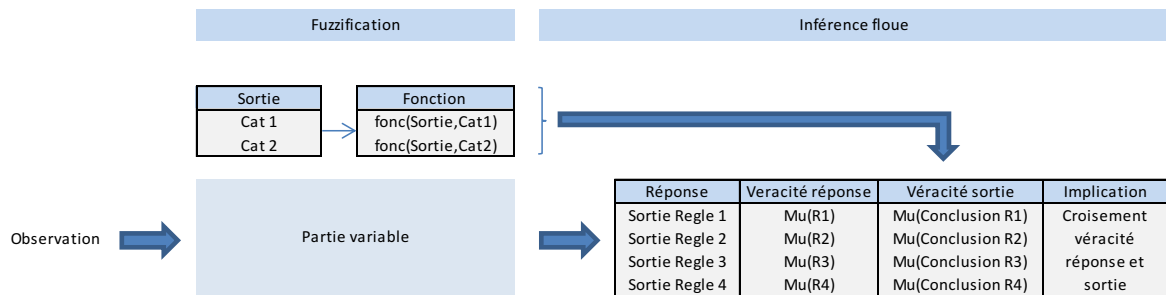
Notons $x_0 = (Variable_1, Variable_2)$ les caractéristiques de l'individu.

- Larsen : $\mu_{Conclusion_{R_i}}' : y \mapsto \mu_{R_i}(x_0) \times \mu_{Conclusion_{R_i}}(y)$
- Mamdani : $\mu_{Conclusion_{R_i}}' : y \mapsto \min_y (\mu_{R_i}(x_0), \mu_{Conclusion_{R_i}}(y))$

Avec :

- $\mu_{R_i}(x_0)$ le degré d'activation de la règle ;
- $\mu_{Conclusion_{R_i}}(y)$ la fonction d'appartenance de l'ensemble flou de sortie en fonction de la règle de décision. Il conviendra de garder en mémoire que $\mu_{Conclusion_{R_i}}$ est une fonction.

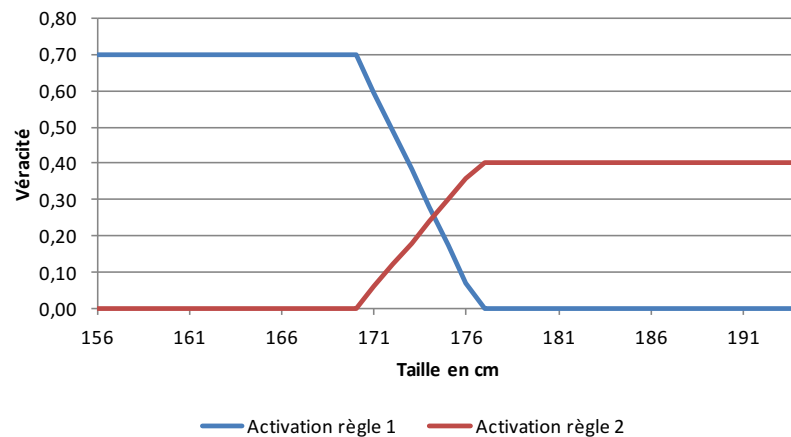
Aussi, calculer la règle d'activation revient à croiser la probabilité de la règle et la probabilité de la sortie associée à la règle. Il est possible de poursuivre l'analogie avec les méthodes bayésiennes du cadre probabiliste classique : la probabilité de la règle peut s'apparenter à une probabilité *a priori*, et la probabilité de la sortie à une probabilité *a posteriori*.



En prenant la règle de Larsen et en reprenant le graphique associé à la règle de sortie et en supposant :

Veracité réponse	Valeur de véracité
Véracité 1	70%
Véracité 2	0%
Véracité 3	40%
Véracité 4	0%

On obtient :



Remarque : L'implication donne naissance à une courbe dans le sens où :

- La variable d'entrée donne naissance à des véracités (en traversant les différentes règles)
- La variable de sortie (associée à chaque règle) est également une fonction de véracité pour chaque modalité.

d. L'agrégation

Cette quatrième étape de l'inférence consiste à regrouper toutes les règles. Ce regroupement est donc effectué à base de « Ou » logiques, ce qui se traduit (cf. Tableau 1) par des « Max ». En reprenant le graphique

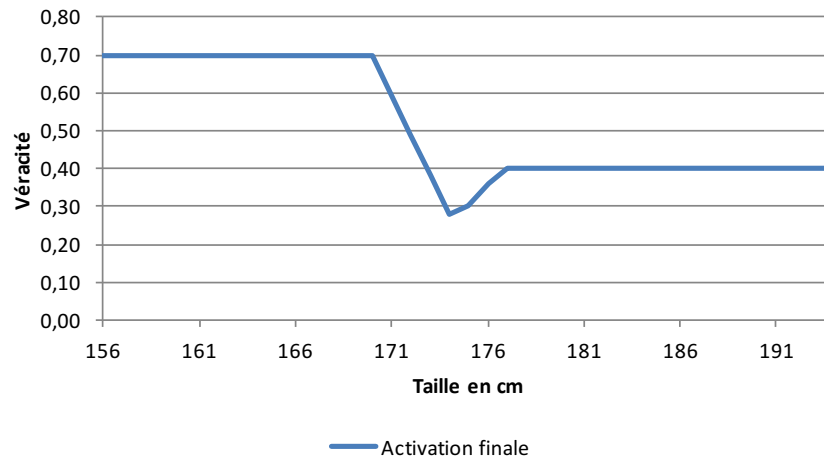


Figure 3 : activation finale

3. La défuzzification

Dernière étape de la logique floue, elle a pour objectif de transformer la courbe d'activation finale obtenue lors de l'étape d'agrégation en une valeur réelle.

Deux méthodes sont alors applicables pour obtenir la valeur retenue de la variable à prédire :

- La méthode de la moyenne des maxima : correspond à la moyenne des valeurs de sortie les plus vraisemblables.
- La méthode des centres de gravité : abscisse du centre de gravité de la surface de la courbe de résultats.

Méthode de la moyenne des maxima

$$MM : x_{MM} = \frac{\int_S x dx}{\int_S dx} = \frac{\sum_{i=0}^n x_i}{N}$$

Où :

$$S = \{x \in U \text{ tel que } \mu(x) = \sup_{x \in U} (\mu(x))\}$$

N le nombre de point appartenant à S .

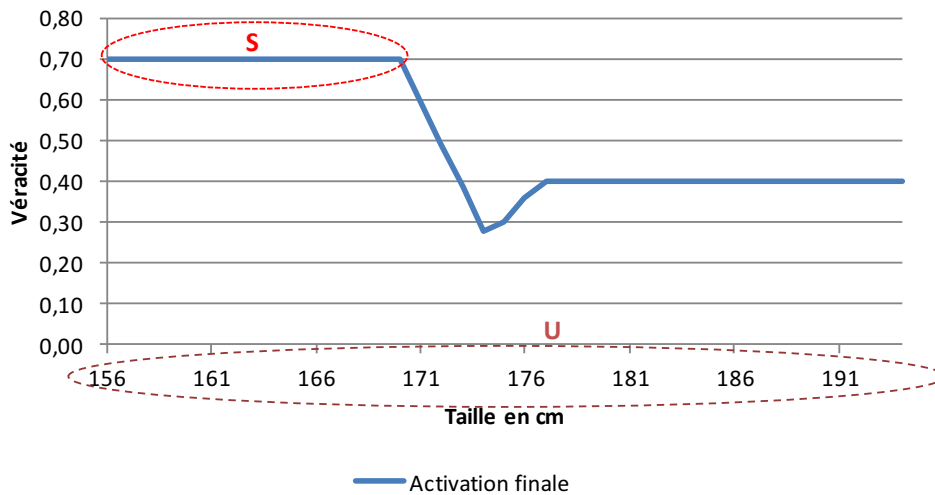


Figure 4 : application méthode des maximas

Cette méthode présente l'inconvénient de fournir des résultats très volatiles (une modification de l'*input* peut conduire à changer la zone de maxima et entraîner une modification brutale de la variable de sortie). Elle est principalement utilisée pour de la reconnaissance de forme.

Méthode des centres de gravité

$$COG : x_G = \frac{\int_U x \mu_x dx}{\int_U \mu_x dx} = \frac{\sum_{i=0}^n x_i \mu_{x_i}}{\sum_{i=0}^n \mu_{x_i}}$$

Où U est l'univers du discours de la variable de sortie. Cela revient à considérer l'espérance liée à la densité $\frac{\mu}{\int_U \mu_x dx}$ associée à la fonction de véracité.

Cette méthode semble préférable (et plus cohérente avec les principes de la logique floue) dans le sens où elle intègre le fait qu'un individu peut appartenir à deux catégories en même temps. On notera l'analogie avec un calcul d'espérance probabiliste (le dénominateur est un facteur de normalisation permettant de prendre en compte le fait que la véracité est supérieure à 100%).

In fine, la logique floue peut se comparer à la statistique classique faisant intervenir des mélanges de lois. Y suit une loi dont la densité dépend de la valeur prise par une variable X de Bernouilli.

$$f_Y = (1 - X)f_{Y|X=0} + Xf_{Y|X=1}$$

$$E[Y] = E_{f_{Y|X=1}}[Y] \times P(X = 1) + E_{f_{Y|X=0}}[Y] \times P(X = 0)$$

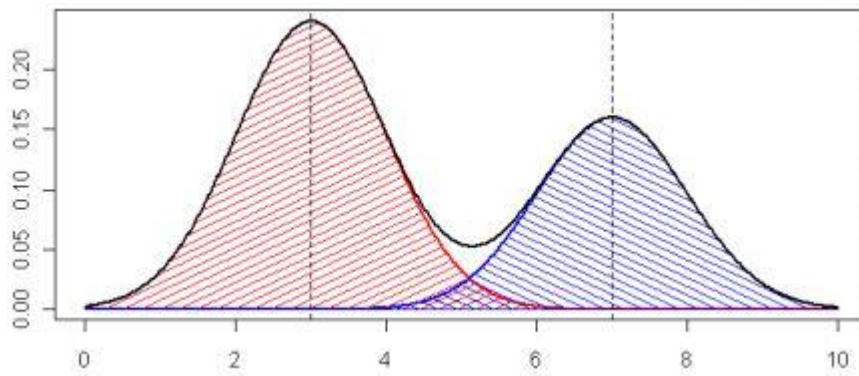


Figure 5 : Exemple de mélanges de loi

4. Conclusion

La logique floue peut s'appliquer dès lors qu'une problématique se base sur une modélisation comportementale ou intègre un raisonnement subjectif (comme par exemple un niveau de satisfaction ou un a-priori). Elle ouvre donc la possibilité à de nombreuses applications actuarielles, notamment :

- la modélisation de comportements associés à un sentiment de satisfaction : rachat conjoncturel, arbitrage Euro/UC.
- la modélisation de décisions liées, en partie, à une intuition : analyse d'un dossier pour sentiment de fraude.

3. EXEMPLE DE PROCESSUS ET D'APPLICATION SUR UN INDIVIDU

1 - Mise en place du processus

Considérons n individus représentés dans une matrice de variables explicatives $\begin{pmatrix} X_{1,1} & \cdots & X_{1,n} \\ \vdots & \ddots & \vdots \\ X_{n,1} & \cdots & X_{n,n} \end{pmatrix}$ et un vecteur de variables réponses : $(Y_1, \dots, Y_n)^t$. Pour l'individu i :

- $X_{i,1}$ = Qualité du résultat
- $X_{i,2}$ = Qualité de la méthode retenue
- $X_{i,3}$ = Qualité de présentation
- Y_i = La note obtenue

Dans un premier temps, on définit les éléments de langage. Par exemple, pour la qualité du résultat :

- **Univers du discours** : note entre 0 et 20
- **Variable linguistique** : qualité du résultat
- **Valeurs linguistiques** : ensemble de modalités regroupant l'ensemble des valeurs qui peuvent être prises par la variable explicative X.

$$R^{Note} = \{\text{"Médiocre"}, \text{"Moyen"}, \text{"Excellent"}\}$$

- $f_{Note}^{R^{Note}}$ une fonction linéaire par morceau, faisant le lien entre Note et R^{Note} :

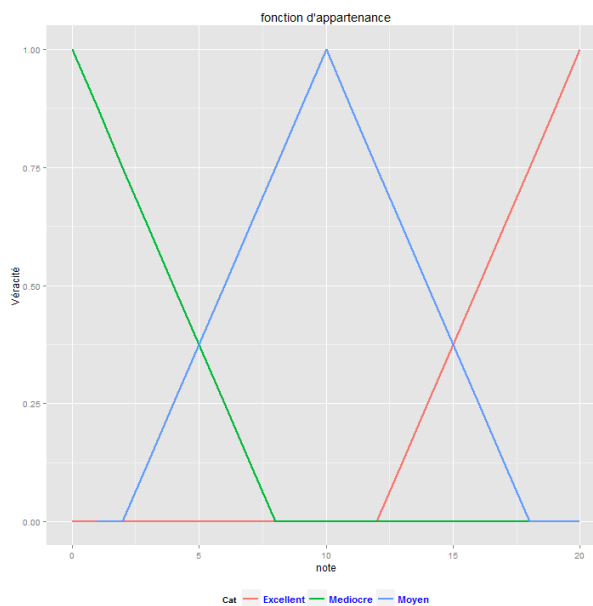


Figure 6 : Fonctions de véracité

2 – L'inférence floue

Une fois les différents éléments du problème traduits en « langage flou », l'inférence a pour objectif de construire des règles de décisions et de trouver pour chacune d'entre elles la règle d'appartenance de la conclusion. La construction de ces règles, principalement basée sur des « ET », des « OU » et des « SANS », se traduit mathématiquement sous la forme ci-après.

Neuf règles sont définies :

- 1- R1 : Si (Résultats = excellent) alors (Evaluation = excellent)
- 2- R2 : Si (Résultats = moyen) alors (Evaluation = moyen)
- 3- R3 : Si (Résultats = médiocre) alors (Evaluation = médiocre)
- 4- R4 : Si (Résultats = moyen) et (Méthodes = médiocre) alors (Evaluation = médiocre)
- 5- R5 : Si (Résultats = moyen) et (Méthodes = excellent) alors (Evaluation = bon)
- 6- R6 : Si (Résultats = médiocre) et (Méthodes = moyen) alors (Evaluation = médiocre)
- 7- R7 : Si (Résultats = excellent) et (Méthodes = excellent) et (Présentation = excellent) alors (Evaluation = excellent)
- 8- R8 : Si (Résultats = médiocre) et (Méthodes = excellent) alors (Evaluation = moyen)
- 9- R9 : Si (Résultats = excellent) et (Méthodes = médiocre) alors (Evaluation = moyen)

La véracité de chaque règle est alors calculée au travers :

- des fonctions d'appartenance,
- des opérateurs flous de Zadeh,
- de la valeur prise par chaque individu.

Par exemple, un individu ayant comme paramétrage Résultats =15, Méthode=16 et Présentation = 16 conduira aux véracités suivantes :

> FResultat(15)

Mediocre	Moyen	Excellent
0	0.375	0.375

> FMethode(16)

Mediocre	Moyen	Excellent
0	0.25	0.5

> FPresentation(16)

Mediocre	Moyen	Excellent
0	0.25	0.5

> Regle[1,]=ActivationRegle(c(15,16,16))

> t(Regle)

R1	0.375
R2	0.375
R3	0.000
R4	0.000
R5	0.375
R6	0.000
R7	0.375
R8	0.000
R9	0.000

Ce stade, l'ensemble des règles ainsi que leur degré d'activation ont toutes été définies. Il convient maintenant de définir l'unique conclusion à apporter à chaque individu. Elle est ici effectuée au travers de la règle de Larsen : $\mu_{Conclusion'_{R_i}}(y) = \mu_{R_i}(x_0) \times \mu_{Conclusion_{R_i}}(y)$

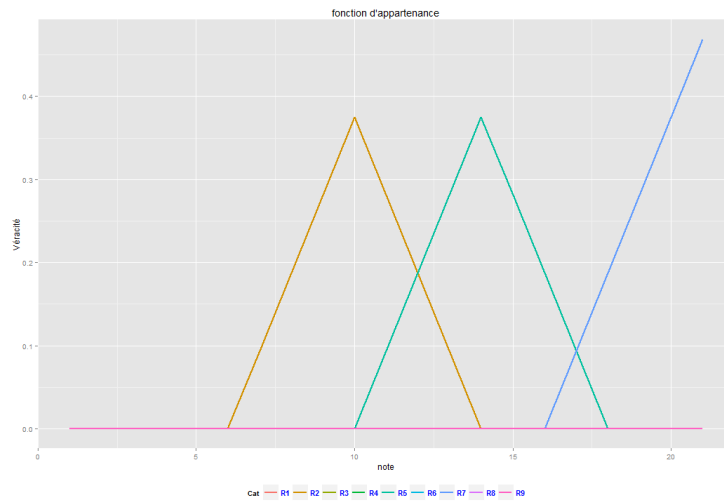


Figure 7 - Larsen pour un input (15,16,16)

L'application de la méthode de Mamdani aurait conduit au graphique suivant :

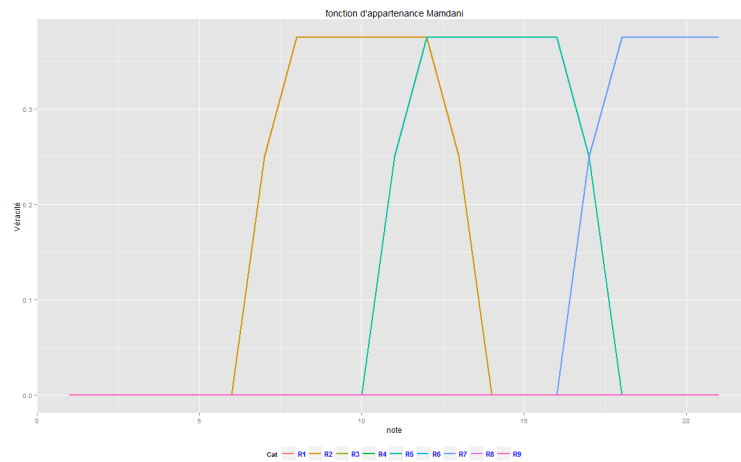


Figure 8 – Mamdani pour un input (15,16,16)

Les différentes vraisemblances (de Larsen) sont ensuite agrégées :

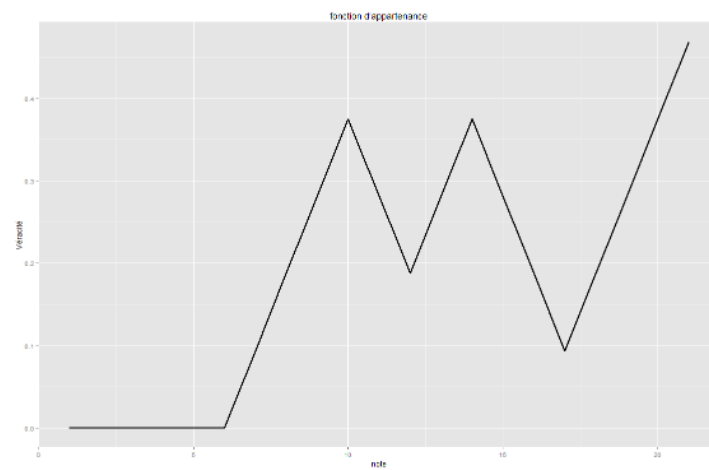


Figure 9-Fonction finale après agrégation pour un input (15,16,16)

3 - La défuzzification

L'application de la méthode de méthode des centres de gravité conduit au graphique suivant, pour une présentation égale à 20 :

