

**PRINCIPAUX DÉTERMINANTS DE LA SINISTRALITÉ EN ASSURANCE
AUTOMOBILE : ÉTUDE DE CAS DE LA SOCIÉTÉ ALGÉRIENNE DES
ASSURANCES (2019)**
**THE MAIN DETERMINANTS OF CLAIMS IN MOTOR INSURANCE: CASE OF
THE ALGERIAN INSURANCE COMPANY (2019)**

***Fadila MEDJTOH eps DIB**

Ecole Nationale de Statistique et d'Economie Appliquée

Email : dibfadila1@yahoo.fr

Zohra AZIBI

Ecole Nationale de Statistique et d'Economie Appliquée

Email : azibizohra@gmail.com

Reçu le: 02/03/2024 **Accepté le :** 11/07/2024 **Publication en ligne le:** 25/07/2024

RESUME:

La sinistralité dans le domaine de l'assurance automobile est évaluée en considérant à la fois la fréquence et le coût des accidents. Cependant, le nombre d'accidents routiers enregistrés annuellement continue d'augmenter, ce qui en fait la huitième cause de mortalité dans le monde selon l'OMS, avec des conséquences significatives en termes de dégâts matériel.

Ce travail vise à analyser les facteurs qui influent sur la sinistralité en assurance automobile, en se basant sur un échantillon de 347 assurés de la Société Algérienne d'Assurance (SAA) pour l'année 2019. Pour ce faire, nous avons utilisé la méthode de régression logistique binaire. Les résultats de notre modèle indiquent que les variables ayant un impact sur la sinistralité comprennent le sexe du conducteur, l'âge du véhicule et son utilisation.

Mots clés: Assurance automobile, sinistralité, régression logistique binaire.

ABSTRACT:

Loss experience in the field of automobile insurance is evaluated by considering both the frequency and the cost of accidents. However, the number of road accidents recorded annually continues to increase, making it the eighth cause of death in the world according to the OMS, with significant consequences in terms of material damage.

This work aims to analyze the factors that influence the loss ratio in automobile insurance, based on a sample of 347 policyholders from the Algerian Insurance Company (SAA) for the year 2019. To do this, we used the method binary logistic regression. The results of our model indicate that variables impacting loss experience include driver gender, vehicle age and vehicle usage.

Keywords: Automobile insurance, claims experience, binary logistic regression.

JEL Classification : G22

*Auteur Correspondant

1. INTRODUCTION :

L'assurance est aujourd'hui un service incontournable qui occupe une place très importante dans la vie de tous les jours et donc dans notre quotidien. Dans les années 1960, elle s'est démocratisée avec notamment le boom automobile. L'assurance se décompose en deux catégories, l'assurance vie et l'assurance non vie. Nous allons nous intéresser à l'assurance non vie en particulier au domaine de l'assurance automobile.

L'assurance automobile constitue l'une des branches essentielles de l'assurance, étant obligatoire pour tout véhicule terrestre à moteur ainsi que toute remorque. Cette assurance couvre non seulement le propriétaire du véhicule, mais aussi toute personne ayant la responsabilité ou la conduite du véhicule, ainsi que les passagers.

Cependant, la branche automobile fait face à un déficit structurel, principalement en raison de la sinistralité élevée qui la caractérise. Cette situation pose des défis aux compagnies d'assurances lorsqu'elles doivent gérer le règlement des sinistres.

La tarification en assurance automobile est un processus complexe visant à évaluer les risques instantanément. La prime d'assurance doit être proportionnelle au niveau de risque encouru par les assurés. Deux types de tarification sont couramment utilisés : la tarification a priori, qui fixe le prix du contrat en fonction des caractéristiques de l'assuré, et la tarification a posteriori, qui ajuste ultérieurement la prime en fonction des antécédents de sinistre, souvent à travers des systèmes comme le "Bonus-Malus".

La sinistralité en assurance automobile est un enjeu majeur, mesuré par la fréquence et le coût des accidents. Dans ce marché hautement concurrentiel, les assureurs cherchent à identifier les facteurs influençant la sinistralité pour mieux segmenter leur portefeuille et établir des classes de risque. Cette démarche vise à garantir une adéquation optimale entre la sinistralité et les primes perçues des assurés.

Les études les plus pertinentes se concentrent sur la relation entre la sinistralité et les caractéristiques de l'assuré, ainsi que sur celles liées au véhicule lui-même.

Parmi les auteurs qui ont entamé ces études on trouve :

(BUHLMANN,1967), (WEDDERBURN et NELDER,1972), (MONFORT et GOURIEROUX,1984), (ZEGER et LIANG,1986), (CF.DIONNE,VANASSE,1989), (STEWART,1994), (CHIAPPORI et CHASSAGNON, 1997,2005), (ORDAZ, MELGAR,et GUERRERO,2006), (VASECHKO, ALBIZZATI, et GRUN-REHOMME, 2009), (BELHOUCHE et NACIB, 2015); (RIAD,2016).

Dans ce travail on s'intéresse aux facteurs explicatifs qui déterminent la sinistralité en assurance automobile (du nombre d'accidents responsables déclarés par l'assuré à son assureur).

Ce travail se concentrera sur la présentation du secteur des assurances en Algérie, avec un accent particulier sur les sinistralités automobiles, on s'intéresse aux facteurs explicatifs

déterminants la sinistralité (l'occurrence des sinistres déclarés par l'assuré à son assureur). **L'objectif de notre travail est d'identifier les variables ayant le plus d'impact sur la sinistralité en assurance automobile en Algérie.**

2. Méthodes et outils :

Notre travail sera traité par la méthode de régression logistique appliquée à un échantillon de 347 assurés de la Société Algérienne d'Assurance (SAA) située à Port Said Alger. Notre objectif sera d'estimer empiriquement la relation entre la probabilité qu'un assuré subisse un accident (1= oui) ou non (0= non), qui constituera notre variable dépendante, et les variables explicatives indépendantes liées à l'assuré et au véhicule.

Pour mener cette recherche, nous devons collecter des données sur les caractéristiques des assurés et des véhicules, telles que le sexe du conducteur, l'âge, l'ancienneté du permis de conduire, l'âge et le type du véhicule, l'usage du véhicule, etc.

En utilisant la méthode de régression logistique, nous serons en mesure d'analyser comment ces variables explicatives influents sur la probabilité d'un accident. Cette approche nous permettra d'obtenir des résultats quantitatifs robustes et d'identifier les facteurs les plus significatifs qui contribuent à la sinistralité en assurance automobile en Algérie. Avant de présenter la méthode et les moyens utilisés nous donnerons un aperçu général sur le marché des assurances en Algérie.

2.1 Le marché de l'assurance automobile en Algérie (2010-2019)

Tableau N°1: l'évolution de chiffre d'affaire (Million de DA) de branche assurance automobile en Algérie

l'année	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
chiffre d'affaire	39644	43552	52446	60922	63879	66202	65253	65504	69 034	69 244
l'évolution	-	0,099	0,2042	0,162	0,049	0,036	-0,014	0,004	0,054	0,003

Source: Réalisation par nous même d'après les données de CNA 2020

Le chiffre d'affaires des assurances automobiles en Algérie a connu une croissance significative au cours de la période allant de 2010 à 2015, passant de 39,64 milliards de dinars en 2010 à 66,20 milliards de dinars en 2015. Cette croissance soutenue est attribuable à divers facteurs, notamment l'augmentation du parc automobile et l'obligation légale d'avoir une assurance automobile.

Cependant, il convient de noter qu'il y a eu un ralentissement dans la croissance du chiffre d'affaires en 2014 et 2015, avec des taux de croissance de 4,9% et 3,6% respectivement. Ce ralentissement peut être attribué à la chute des importations algériennes de véhicules au cours de cette période, ce qui a réduit le nombre de nouveaux véhicules nécessitant une assurance.

PRINCIPAUX DÉTERMINANTS DE LA SINISTRALITÉ EN ASSURANCE AUTOMOBILE : ÉTUDE DE CAS DE LA SOCIÉTÉ ALGÉRIENNE DES ASSURANCES (2019)

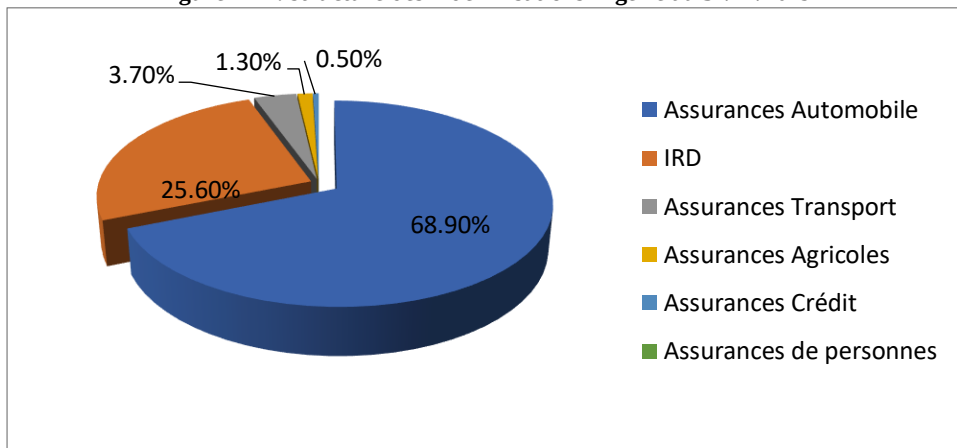
En outre, en 2016, le chiffre d'affaires des assurances automobiles a connu une baisse de 1,51% avec un chiffre d'affaire de 65.25 milliards. Cette baisse peut être liée à divers facteurs économiques et réglementaires, ainsi qu'à des conditions de marché changeantes. Exemple les prix des véhicules qu'ont fortement augmentés dans cette période. En 2019 il passe à 69.244 milliards de dinar avec un taux annuel moyen de 0.3%.

2.2 Situation des sinistres des assurances de dommages au 31/12/2019

Les sinistres des assurances de dommages affichent une augmentation de 20.5% avec un montant de 77.5 milliards de DA par rapport à 2018. Cette hausse tirée par les branches « automobile » et « crédit », avec des taux d'évolution respectifs de 16.8 % et 20.05% .la branche transport marque une régression de 34,8%.

2.2.1 Structure des indemnisations:

Figure N°1 : structure des indemnisations Algérie au 31/12/2019



Source : Réalisation par nous même d'après les données de CNA

Les indemnisations versées par les sociétés d'assurances de dommages atteignent au 31/12/2019, un montant de 68,7 milliards de DA. La part des règlements des dossiers sinistres de la branche « automobile » est de 68.9% pour un montant de 47,4 milliards de DA. La branche « IRD » avec un taux de 25%.

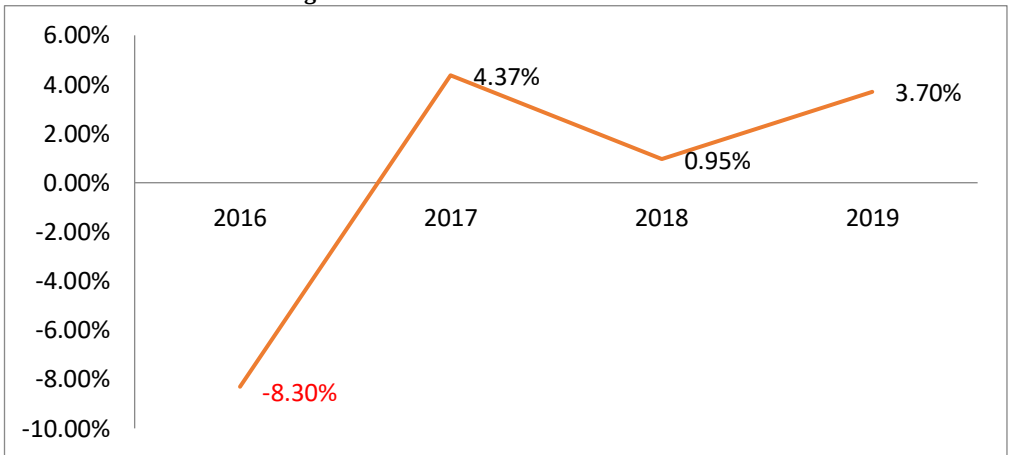
2.2.2 la sinistralité automobile en Algérie

Tableau N°2: l'évolution des accidents de la route en Algérie.

l'année	2015	2016	2017	2018	2019
Indemnisations en DA	47 327 242 478	43 396 955 475	45294 353 601	45 725 979 100	47 418 052 702
évolution	-	-8,30%	4,37%	0,95%	3,70%
sinistres à payer en DA	34 657 265 287	35 179 393 282	35110 461 894	35 339 822 947	35 797 785 197
évolution %	-	1,51%	-0,20%	0,65%	1,30%

Source : Réalisation par nous même d'après les données de CNA 2020

Figure N°2 : taux d'évolution des Indemnisations

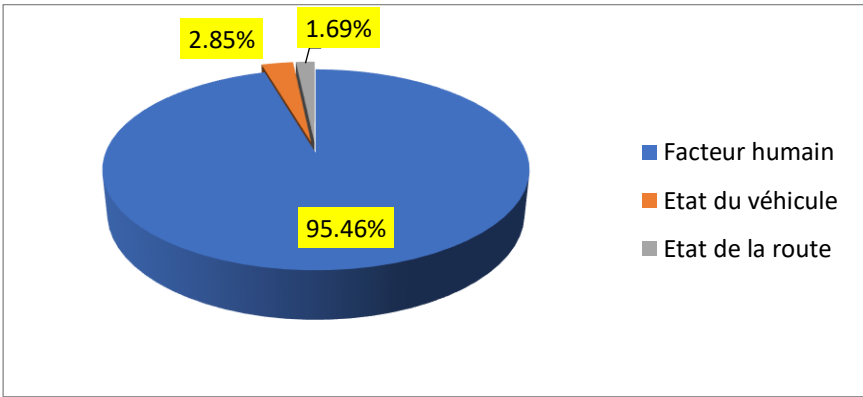


Source : Réalisation par nous même d'après les données de CNA2020

Les indemnisations versées par la branche assurance « automobile » atteignent 47.3 milliards de DA en 2015. Et affichent une régression de 8.3% en 2016, Ensuite, ces chiffres augmentent pour atteindre 45,29 milliards de DA en 2017, représentant une hausse de 4,37 % puis de 0.95% en 2018 et atteignent 47.41 milliards de DA en 2019 par un taux d'évolution de 3.70% par rapport à l'année 2018.

D'après le tableau N°2, les sinistres à payer par la branche des assurances automobile s'élèvent à 35,79 milliards de dinars pour l'année 2019, comparativement à 34,65 milliards de dinars en 2015, ce qui représente une augmentation de 3,29 %.

Figure N°03 : Les éléments des accidents de la route en Algérie



Source :Focus sur la sécurité routière. 2016

Les accidents sont principalement causés par des facteurs humains, représentant 95,46 % des cas, tandis que l'état du véhicule contribue à 2,85 % des accidents et l'état de la route à 1,69 %.

2.2.3 Évolution des sinistres de la route

Tableau N°3: Nombre de blessés, de décédés , d'accidents de la route 2011-2019

année	2011	2012	2013	2014	2015	2016	2017	2018	2019
Fréquence d'Accident	42788	49426	56136	64294	64813	59902	57328	54241	53802
Nombre de Blessés	50958	61226	68734	80225	79828	73337	69754	65959	65353
Nombre de décédées	2293	2222	2294	2519	2332	2118	1882	1779	1564

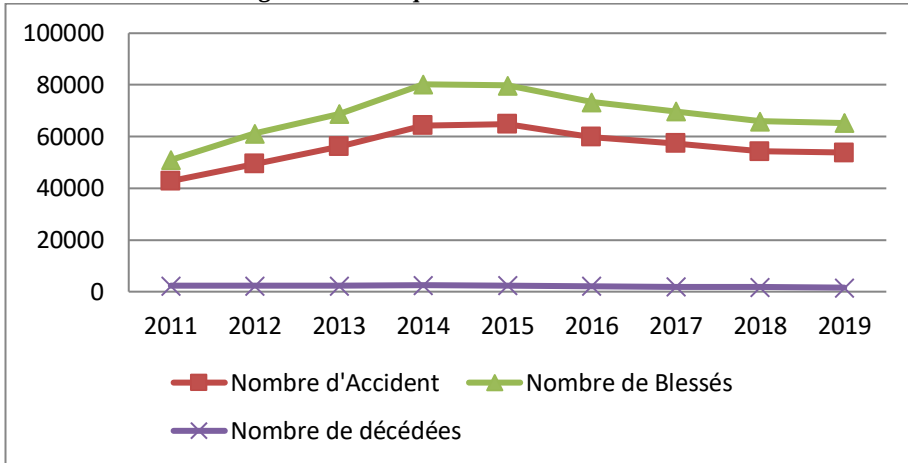
Source : Réalisation par nous même d'après les données de protection civile (Civile, 2023).

Les données présentées dans le tableau indiquent que pendant la période de 2011 à 2015 est varié de 42 788 à 64813 accident avec un taux d'évolution de 51.47%, cette augmentation cause une augmentation du nombre des blessés de 50958 à 80225 blessé en 2014 et un faible régression a 79828 blessés. Les décédées connue une faible augmentation dans cette période avec un taux de 1.7%.

De 2015 à 2019, on observe une diminution du nombre d'accidents de 16,98 % par rapport à 2015. De même, le nombre de blessés a également connu une baisse de 18,13 %, tandis que le nombre de décès a diminué de 32,93 %. Cette réduction peut être expliquée par les efforts des autorités publiques, qui ont intensifié les actions et les initiatives au cours de ces dernières années. Ces efforts comprennent le durcissement des sanctions du code de la route, la révision des programmes de formation des conducteurs, l'amélioration de l'état du réseau routier, la généralisation du contrôle technique des véhicules, ainsi que la

modernisation des équipements de contrôle des services de sécurité. Ces différentes mesures ont contribué, dans une certaine mesure, à réduire l'insécurité routière.

Figure N°4 : Fréquence des sinistres de la route 2011-2019



Source : travail fait par nous même d'après les données de la protection civile

3. Aspects théorique de modèle linéaire généralisé

3.1 Description du modèle:

Le modèle linéaire généralisé (GLM) (Tufféry, 2012, pp. 340-348) est apparu (NELDER et WEDDERBURN 1972). C'est un modèle adapté à de nombreuses problématiques et plus particulièrement dans le domaine de la statistique.

Le modèle de GLM a un avantage par rapport aux modèles linéaires: le caractère normal des variables à expliquer n'est plus imposé, seule une appartenance à une famille exponentielle est indispensable (Planchet, Thérond, & Jacquemin, 2005),

3.2 Les éléments du GLM:

En effet Le modèle linéaire généralisé se compose de trois éléments principaux : la fonction de lien, la composante aléatoire et la composante déterministe.

3.2.1 La Composante aléatoire :

Dans le cas où la variable à expliquer, notée "Y", correspond au nombre de sinistres d'un véhicule et que ses densités appartiennent à la famille exponentielle, la composante aléatoire du modèle linéaire généralisé (GLM) est représentée par la distribution de la variable "Y". Cette distribution peut être choisie en fonction des caractéristiques des données, telles que la variance et la structure des résidus.

3.2.2 L'élément déterministe

L'ensemble des variables explicatives, représentées par " X_i ", est utilisé pour déterminer les variables expliquées, notamment le nombre de sinistres. Dans cette étude, la

composante déterministe est formée de six variables spécifiques, qui seront détaillées dans la section suivante.

3.3 La régression logistique : un cas particulier de GLM

Nous considérons une variable binaire Y , qui peut prendre deux valeurs ($Y = 0$ ou $Y = 1$). L'objectif est d'expliquer Y en fonction d'un vecteur composé de p variables explicatives. $X = (X^1, \dots, X^p)$ $\pi(X) = p(Y = 1/X)$ $Y = \pi(X) + \varepsilon$

L'erreur $\varepsilon = Y - \pi(X)$ peut prendre seulement deux valeurs, à savoir :

$$\varepsilon = \begin{cases} 1 - \pi(X) & \text{si } Y = 1 \text{ avec la probabilité } \pi(x) \\ -\pi(x) & \text{si } y = 0 \text{ avec la probabilité } 1 - \pi(x) \end{cases}$$

La régression logistique peut être envisagée comme un cas spécifique d'un modèle plus vaste dans lequel $\pi(x)$ est représenté selon la forme suivante :

$$\pi(x) = G(\beta_0 + \beta_1 x_0 + \dots + \beta_p x_p)$$

Avec

Dans ce contexte, G est une fonction connue, croissante, qui prend ses valeurs dans l'intervalle $[0,1]$.

Les paramètres de la régression logistique à estimer sont représentés par (β_j) pour $0 \leq j \leq p$, où p est le nombre de variables explicatives.

Le modèle logistique correspond à une situation où :

$$G(u) = \frac{e^u}{1 + e^u}$$

4. Définition et analyse des variables

4.1 Analyse de l'enchantions et le modèle:

Dans cette section, nous présentons l'échantillon utilisé ainsi que le modèle sous-jacent à notre analyse. Notre échantillon se compose de 347 individus, représentant un échantillon d'assurés sous contrat d'assurance automobile particulier auprès de l'Agence d'assurance "saa-1501-PORT SAID" sur une période allant du 1er janvier 2019 au 31 septembre 2019. Ce groupe a été constitué de manière aléatoire et servira de base à notre analyse. En se basant sur la disponibilité des données et la littérature existante sur les accidents de la circulation, nous avons sélectionné un ensemble de variables pour une segmentation préalable. Ces variables permettent de classer les assurés en différents groupes, que nous appelons les groupes de risque.

Nous disposons de trois ensembles de variables pour chaque assuré de notre échantillon : les caractéristiques du conducteur, celles du véhicule, et les données sur les sinistres. Pour être plus précis :

4.2 Codage des variables

4.2.1 Attributs du conducteur :

Sexe de conducteur (*sexe*) : Cela concerne le genre du conducteur principal déclaré..

$$sexe = \begin{cases} 1. Homme \\ 2. femme \end{cases}$$

L'âge de conducteur (*âge_cond*) : Exprimer en années.

$$âge_cond = \begin{cases} 1 & âge_cond \leq 25 \\ 2 & 25 < âge_cond \leq 40 \\ 3 & âge_cond > 40 \end{cases}$$

L'ancienneté de permis de conduite de conducteur (*âge_permis*)

$$âge_permis = \begin{cases} 1 & âge_permis \leq 10 \\ 2 & 10 < âge_permis \leq 20 \\ 3 & âge_permis > 20 \end{cases}$$

4.2.2 Attributs du véhicule :

Ancienneté (durée de mise en marche de véhicule) (*âge_veh*):

$$âge_véh = \begin{cases} 1 & 0 < âge_vih \leq 10 \\ 2 & âge_vih > 10 \end{cases}$$

La manière dont le véhicule est utilisé (*usg*) : (*usg*)

$$usg = \begin{cases} 1 & Fonctonnaire \\ 2 & Affaire \\ 3 & Commerce \\ 4 & Taxi \end{cases}$$

La puissance du véhicule (*puis*)

$$puis = \begin{cases} 1 & puis \leq 5 \\ 2 & 5 < puis \leq 10 \\ 3 & puis > 10 \end{cases}$$

4.2.3 Détails sur les sinistres :

Nombre de réclamations déclarées (nbr_s) : pour l'année de référence, ici 2019.

Fréquence des sinistres ($oc.s$) : créé à partir de fréquence de sinistres ($nbr.s$)

$$Oc.s = \begin{cases} 0 & \text{si } nbr.s = 0 \\ 1 & \text{si } nbr.s > 0 \end{cases}$$

4.3 Analyse uni-varié :

4.3.1 Nombre de sinistres déclaré :

Tableau N°4: Distribution de la fréquence des sinistres

Classes	Nbr .s	%	% cumulé
0	233	67.15%	67,15%
1	89	25.65%	92,80%
2	16	4.61%	97,41%
3	7	2.01%	99,42%
4	1	0.29%	99,71%
5	1	0.29%	100,00%
Totale	347	100%	100,00%

Source : Réalisation par nous même à partir du base de données. SAA2019.

On constat que 233 assurés ne déclare aucun sinistre avec un pourcentage de 67,15%, 25.65% ont réalisé un seul sinistre pendant l'année 2019, 7.20% des conducteurs assurés réalisent plus d'un sinistre donc, on peut dire que la réalisation des sinistre est un événement rare.

4.3.2 Le Sexe (genre) du conducteur

Tableau n°5 : Distribution des assurés en fonction du sexe et du type de conducteur

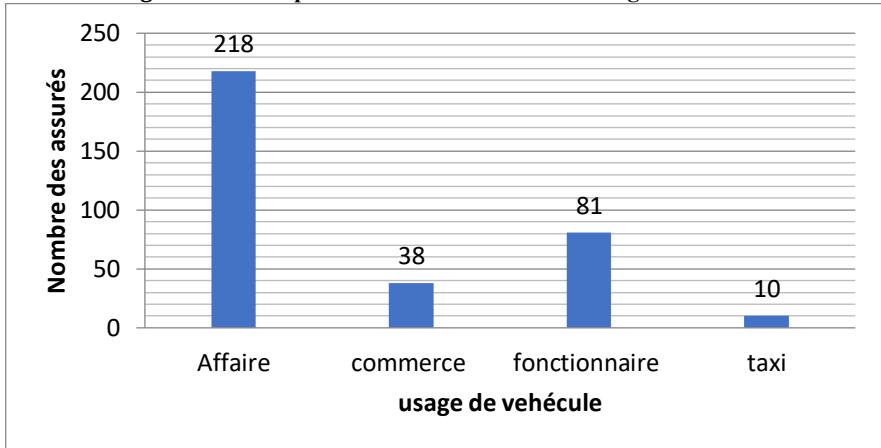
	Effectif	%
Sexe féminin	72	20.75%
Sexe masculin	275	79.25%
Total	347	100%

Source : Réalisation par nous même à partir de la base de données SAA2019.

On constate que cette échantillon constitue de 275 assurés de sexe masculin avec un pourcentage de 79.25%, par contre les femmes représentent 20.75% (72 femmes).

4.3.3 L'usage du véhicule

Figure n°5 : la répartition des assurés selon L'usage du véhicule

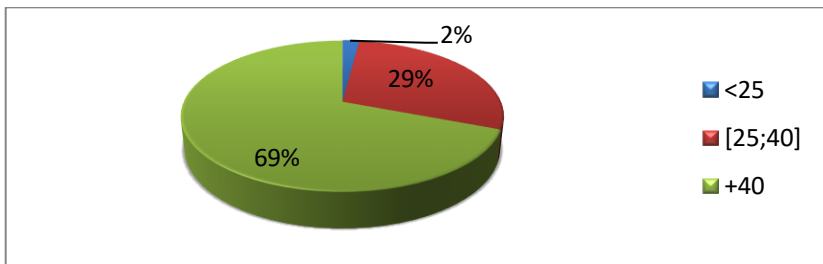


Source : Réalisation par nous même à partir de la base de données SAA2019.

Nous apprenons que près de 218 des véhicules ont souscrit à l'usage « affaire » avec un pourcentage de 62,82 %, 38 véhicules à l'usage «commerce », 81 et 10 véhicules respectivement pour l'usage « fonctionnaire » et « taxi ».

4.3.4 L'âge de conducteur

Figure N°6 : la répartition des assurés en fonction de l'âge de conducteur

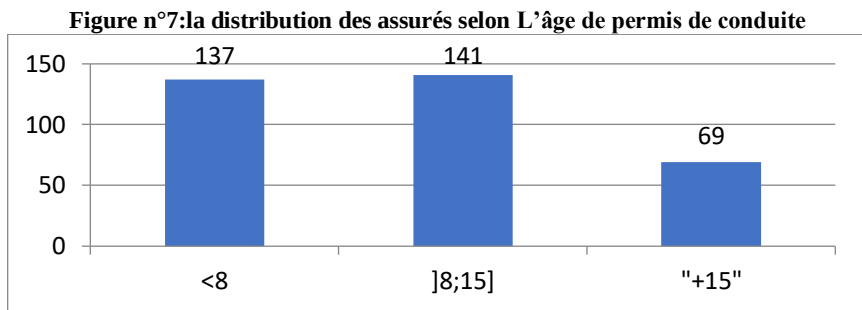


Source : exploration à partir des données de la SAA2019.

On constate que 69% des conducteurs assurés sont de l'âge supérieurs aux 40 ans, 29% des assurés sont entre 25 ans et 40 ans et 2% sont inférieure à 25 ans .

4.3.5 L'âge de permis de conduite

**PRINCIPAUX DÉTERMINANTS DE LA SINISTRALITÉ EN ASSURANCE
AUTOMOBILE : ÉTUDE DE CAS DE LA SOCIÉTÉ ALGÉRIENNE DES
ASSURANCES (2019)**



Source : exploration à partir des données de la SAA2019.

On constate que 137 conducteur ont un âge de permis de moins de 8 ans qui représente 39,48% de l'échantillon, 40,36% ont un âge de permis entre 8 et 15 ans, finalement, 19,88 % des assurés ont un permis de conduire plus de 15 ans.

4.3.6 Variables continus

Tableau n°7 : Statistiques descriptives des variables quantitatives utilisées

	Moyenne	Médiane	Mode	Écart-type	Minimum	Maximum
âge du véhicule	7,67435159	6	5	6,55748298	0	56

Source : à partir de la base de données SAA2019 obtenue à l'aide du logiciel Excel 2010.

4.4 Analyse bi-varie

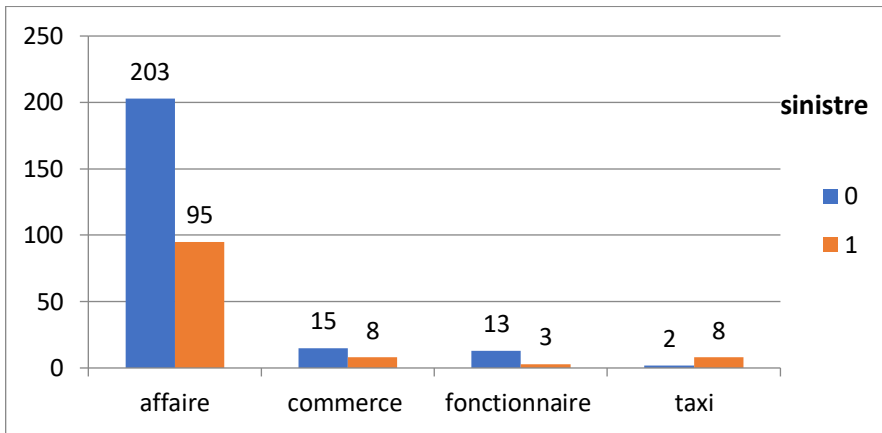
Tableau n°8 :La distribution des sinistres en fonction du sexe des assurés

		Ne présente pas de sinistre	Présente de sinistre	Total
Assuré	Femme	33	39	72
	Homme	200	75	275
Total		233	114	347

Source : à partir de la base de données SAA2019 obtenue à l'aide du logiciel Excel 2010.

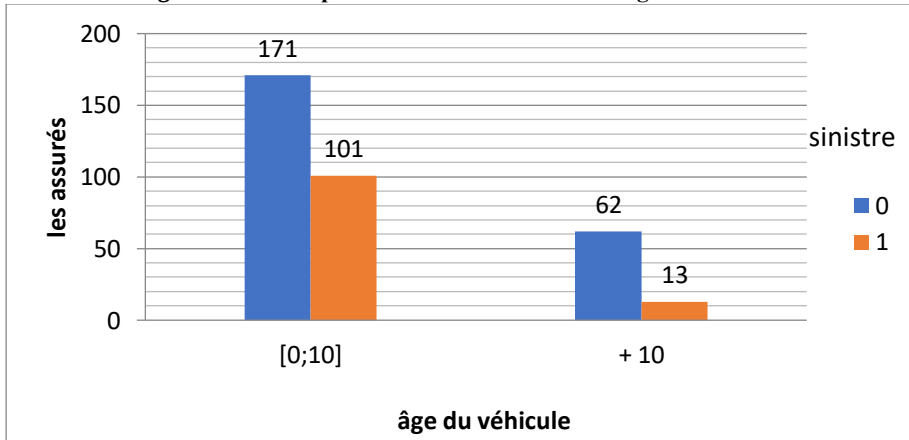
Le tableau montre une sous-sinistralité pour les hommes (sur 275 homme on a 75 réalisé au moins un sinistre qui équivalent au 27.3% de totale des hommes) contrairement aux femmes (54.2 % des femmes en réalisé un sinistre durant l'année 2019)

On constate concernant la sinistralité en terme générale que 32.85% des assurés ont déclaré au moins un sinistre qui équivalent au 114 assuré avec sinistre contre 233 sans sinistre .

Figure n°8: La distribution des sinistres selon l'utilisation du véhicule

Source : exploration à partir de la base de données SAA2019.

On constate que 83,33% des véhicules assurés véhicules à usage «Affaire» ont causé au moins un sinistre, un pourcentage supérieure aux autres usages (commerce, fonctionnaire, taxi) Qui sont respectivement 7,01%, 2,63% et 7.01% on constate que les fonctionnaires sont les conducteurs les plus prudents.

Figure n°9 : La répartition des sinistres selon l'âge du véhicule

Source : Réalisation par nous même à partir de la base de données SAA2019.

On observe que la fréquence des sinistres diminue avec l'ancienneté du véhicule. En effet, les conducteurs de véhicules âgés de moins de 10 ans sont responsables de plus d'accidents que les autres classes d'âge, avec 88,60 % de sinistres déclarés, tandis que les véhicules âgés de plus de 10 ans ont moins d'incidents, avec un taux de sinistralité de 11,40 %. Cela peut s'expliquer par le fait que les conducteurs de véhicules plus anciens sont généralement plus prudents.

PRINCIPAUX DÉTERMINANTS DE LA SINISTRALITÉ EN ASSURANCE AUTOMOBILE : ÉTUDE DE CAS DE LA SOCIÉTÉ ALGÉRIENNE DES ASSURANCES (2019)

5. Essai d’analyser les factures déterminant de la sinistralité en assurance automobile

Nous allons utiliser l'approche d'apprentissage automatique en utilisant la régression logistique binaire sur R pour modéliser l'occurrence des sinistres en fonction de variables explicatives liées au facteur humain (sexe, âge du conducteur, ancienneté du permis) et au véhicule (usage, âge du véhicule, puissance). Nous chercherons à prédire la probabilité qu'un assuré ait un sinistre (1=oui) ou non (0=non) pour l'ensemble de la population.

La variable dépendant $Y= oc_s$ occurrence des sinistres

5.1 Etape 1 la division base de donné

Pour suivre l’approche d’apprentissage automatique, le jeu de données se divise comme suit.

Ensemble d’entraînement (train) : 70 % de l'ensemble complet des données est réservé pour construire le modèle.

Ensemble de teste (test) : 30% réservé pour l’évaluation de modèle estimé a partir des donné (train)

5.2 Etape 2 : Phase d’apprentissage

Dans cette phase on va construire le modèle à partir des donnés entraînement (train)

5.2.1 Modèle initiale

Une fois que toutes les variables explicatives ont été intégrées dans le modèle, on obtient :

Tableau n°10 : modèle initiale

Coefficients :					
	Estimate	Std.Error	z value	Pr(> z)	
(Intercept)	-5.92215	1.9325	-3.966	7.31e-06	***
sexe2	2.46399	0.51739	4.762	1.91e-06	***
âge_cond	0.01879	0.01506	1.248	0 ,212170	
âge_ permis	-0.02639	0 ,02593	-1,017	0 ,308922	
puis	0,10942	0,10150	1,078	0,281055	
usg2	4.21362	1.09475	3.849	0,000119	***
usg3	4.97331	1.18347	5.047	4.48e-07	**
usg4	4,98898	1,56323	4,471	7 ,79 e -06	**
âge_veh	-0.09365	0.04065	-2.304	0,21241	**
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Source : à partir de la base de données SAA2019 obtenue à l’aide du logiciel R

Nous observons que les coefficients des variables ne sont pas tous significatifs. Cela concerne spécifiquement les coefficients associés à l'ancienneté du permis de conduire (« âge_permis ») à l'âge du conducteur (« âge_cond ») et à la puissance du véhicule (« puis »).

5.2.2 la démarche de choix des variables

La procédure de sélection des variables descendante, également appelée Backward Selection, consiste à calculer l'AIC du modèle complet (avec toutes les variables explicatives disponibles) ainsi que l'impact du retrait de chaque variable sur l'AIC. Ensuite, le processus élimine du modèle la variable dont le retrait permet une baisse significative de l'AIC.

Tableau n°11: choix des variables

Start: AIC=226.88		
oc_s ~ sexe + âge_cond + âge_permis + usg + puis + âge_veh		
DfDevianceAIC		
- âge_permis1	209.93	225.93
- puis 1	210.04	226.04
- âge_cond 1	210.46	226.46
<none>	208.88	226.88
- âge_veh 1	216.39	232.39
- sexe 1	237.85	253.85
- usg 3	280.84	292.84
		⋮
Step: AIC=224.38		
oc_s ~ sexe + âge_cond + usg + âge_veh		
DfDeviance AIC		
<none>	212.38	224.38
- âge_veh1	219.37	229.37
- sexe1	240.54	250.54
-usg 3	284.02	290.02

Source :traitement par logiciel R à partir de la base de données SAA2019

Suite à l'application de la procédure descendante, les variables *âge_permis* et *puiss*, ne font plus partie du modèle.

Nous remarquons que le coefficient du variable *âge_cond* n'est pas significatif donc on rejette cette variable du modèle on refait l'estimation du modèle.

On obtient les résultats suivant

**PRINCIPAUX DÉTERMINANTS DE LA SINISTRALITÉ EN ASSURANCE
AUTOMOBILE : ÉTUDE DE CAS DE LA SOCIÉTÉ ALGÉRIENNE DES
ASSURANCES (2019)**

Tableau n°12: modèle finale

	Coefficients	Std.Error	z value	Pr(> z)	
(Intercept)	-5.6096	1.13154	-4.074	4.62e-06	***
sexe2	2.3672	0.5012	4.723	2.33e-06	***
usg2	4.2475	1.08449	3.843	0,000121	***
usg3	4.5890	1.16420	5.020	5.16e-07	***
usg4	4.7334	1.15450	4.383	1.17e-05	* ***
âge_veh 2	-0.08937	0.3984	-2.244	0.02484	* *
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Source : à partir de la base de données SAA2019 obtenue à l'aide du logiciel R

On constate que tous les variables explicatives représente des coefficients significatives au seuil de 1% .Les variable : sexe de conducteur, âge du véhicule et l'usage ont un effet significative sur la sinistralité automobile. On a retenus Trois variables et la fonction de lien utilisée est la fonction logit, on a alors.

$$\ln \left(\frac{p(Y = 1/X)}{P(y = 0/X)} \right) = -4.6096 + 2.3672 \times \text{sexe2} + 4.2475 \times \text{usg2} + 4.5890 \times \text{usg3} + 4.7334 \times \text{usg4} - 0.08937 \times \text{age.veh2}$$

Les coefficients du modèle logistique, notés β_j , sont souvent interprétés en termes d'odds ratio.

$$\text{odds}(x_j) = e^{\beta_j}$$

Tableau n°13 :Coefficients et Odds du modèle

	β_i	$Odds(x_i) = \exp(\beta_i)$
(Intercept)	-5.6096	0.003662534
sexe2	2.3672	10.66748148
usg2	4.2475	69.93036771
usg3	4.5890	98.39598496
usg4	4.7334	113.6814229
âge_veh2	-0.08937	0.914507143

Source : traitement par logiciel R à partir de la base de données SAA2019

Le logit de la constante indique que pour un assuré de sexe masculin (sex=1) et utilisant son véhicule à des fins professionnelles (usage=1, "fonctionnaire") avec un véhicule âgé de moins de 10 ans, le odds est de 0.0036662. Cela signifie que le modèle prédit que

seulement 0.36 % des hommes assurés, utilisant leur véhicule à des fins professionnelles et ayant un véhicule de moins de 10 ans, auront un sinistre.

Pour les coefficients des variable On prend par exemple variable sexe on a *odds* (pour sexe2)=10.667 le risque de sinistralité augmente avec le sexe féminin jusqu'à être multiplié par 10.66.

Quand l'âge de véhicule augmente d'une unité (année le risque d'avoir un sinistre sera multiplié par 0.9145 (diminué de 8.54%).

5.2.4 Qualité d'ajustement

Tableau n°14 :Le résidu de déviance

Le résidu de déviance	$q_{n-p-1}^{1-\alpha}$
212.39	274.99

Le modèle à un résidu de déviance égale à 212.39 D'après le tableau ci-dessus nous pouvons conclure que la déviance observée est inférieure au quantile à 95% d'ordre 238 (n-p-1), avec n: notre échantillon, p:valeur de probabilité(le paramètre de la regression a estimer) on peut donc admettre que le modèle est pertinent.

5.3 Etape 3 : phase d'évaluation de modèle

Dans cette phase on va utiliser le modèle qu'on a estimé pour prédire des classes des assurées avec sinistre ou sans sinistre de la base de donnée « teste ». On prend la classe « avec sinistre » comme un classe positif et « sans sinistre » négatif

On obtient les résultats suivants :

5.3.1 Matrice de confusion

Tableau n°15 : Matrice de confusion

Observation	Préversion	
	Non (-)	Oui (+)
Non (-)	64	13
Oui (+)	5	21

Source : à partir de la base de données SAA2019 obtenue à l'aide du logiciel R

D'après la matrice de confusion on remarque que :

- Le nombre d'assurés observé comme « sans sinistre » et que le modèle a bien classé est de 64 assurés

- Le nombre d'assurés observé comme sans sinistre et que le modèle a classé comme « avec sinistre » est de 13 assurés

**PRINCIPAUX DÉTERMINANTS DE LA SINISTRALITÉ EN ASSURANCE
AUTOMOBILE : ÉTUDE DE CAS DE LA SOCIÉTÉ ALGÉRIENNE DES
ASSURANCES (2019)**

- Le nombre d'assuré que le modèle a bien classé comme « avec sinistre » est de 21 assurées.

- Le nombre d'assuré observé « avec sinistre » et que le modèle a prédit comme « sans sinistre » est de 5 assurés.

5.3.2 Mesures de performances

Les plus importants indicateurs synthétiques qui ont été calculés à partir de la matrice de confusion sont :

Tableau n°16 : indicateurs de mesure de performances de modèle

Indicateur	La Spécificité	Le Rappel	Accuracy	F-Mesure	Précision
Formule ¹	$\frac{VN}{VN+FP}$	$\frac{VP}{VP+FN}$	$\frac{VP+VN}{P+N}$	$\frac{2*(précision+rappel)}{précision+rappel}$	$\frac{VP}{VP+FP}$
Valeur	92.75%	61.76	82.52%	74.14%	86.36%

- « accuracy » : Le modèle a bien classé 82.52% des assurés

- « Taux d'erreur » : le modèle a trompé dans la classification de 17.48% des individus

- « La Spécificité »: permis les assuré prévue par le modèle comme « sans sinistre » , 92.75% sont bien classé.

- «Le Rappel » : 61.76% des assurés observé « avec sinistre », sont bien classé par le modèle.

- Le modèle à réussi de classé 68.36% parmi les assuré qui on déclarer un sinistre.

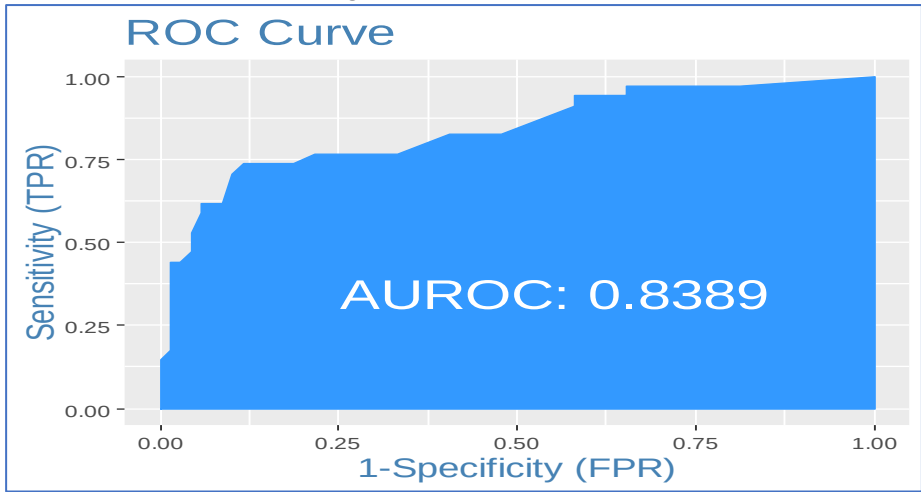
¹ VP : Vraies positives

FP : Fausses positives

FN :Fausses négatives

VN :Vraies négatives

Figure N°10 : Courbe ROC



Source :traitement par logiciel R à partir de la base de données SAA2019

Avec une valeur de l'AUROC (Area under Receiver Operating Characteristic), égale à 83.89% et f-mesure égale 70.54%,et une précision de 86.36% et un taux d'erreur de 17.48 % le modèle permis un bon classement des individus

Nous avons établi un modèle de régression logistique pour prédire l'occurrence des sinistres en utilisant des variables telles que le sexe de l'assuré et des caractéristiques liées à son véhicule, notamment la classe d'âge du véhicule et la catégorie d'usage. Les résultats de validation du modèle sont concluants, ce qui indique que le modèle est efficace pour prédire les sinistres avec précision.

Les résultats empiriques nous montrent que les facteur les plus influent sur la sinistralité automobile sont l'usage du véhicule, sexe de conducteur, et l'ancienneté du véhicule.

CONCLUSION GENERALE

L'assurance automobile est un pilier de la protection contre les aléas de la vie, répondant au besoin inné de sécurité tout en servant d'outil de prévoyance et de complément essentiel aux transactions juridiques. Avec plus de 6,4 millions de véhicules assurés en 2018 en Algérie, représentant le secteur le plus important de l'assurance non-vie avec un chiffre d'affaires de 69,24 milliards de dinars en 2019, l'assurance automobile est une activité d'un grand intérêt public.

Cependant, cette branche est confrontée à une sinistralité élevée, avec 47,41 milliards de dinars en indemnisation enregistrés en 2019, posant des défis pour les compagnies d'assurance dans le processus de règlement des sinistres.

PRINCIPAUX DÉTERMINANTS DE LA SINISTRALITÉ EN ASSURANCE AUTOMOBILE : ÉTUDE DE CAS DE LA SOCIÉTÉ ALGÉRIENNE DES ASSURANCES (2019)

Dans ce contexte concurrentiel, les compagnies d'assurance cherchent à identifier les facteurs qui influent sur la sinistralité automobile. Notre étude empirique, basée sur un échantillon aléatoire de 347 assurés de l'agence SAA agence 1501, a permis de confirmer que la sinistralité automobile est influencée par des facteurs liés à l'assuré (tel que le sexe) et au véhicule (comme l'ancienneté et l'usage).

Nous avons constaté que la sinistralité diminue avec l'ancienneté du véhicule, tandis que les femmes sont plus susceptibles d'être impliquées dans des sinistres que les hommes. Par ailleurs, une sur-sinistralité a été observée chez les véhicules utilisés à des fins Affaires .

Cependant, notre étude présente des limites, notamment la taille réduite de l'échantillon et le fait que d'autres facteurs externes tels que l'état des routes et la signalisation n'ont pas été pris en compte. En perspective, une étude plus vaste, incluant un échantillon plus représentatif de plusieurs compagnies, pourrait être envisagée. De plus, l'analyse pourrait intégrer d'autres aspects environnementaux, permettant de mieux comprendre l'impact relatif des facteurs humains sur la sinistralité automobile.

Bibliographique:

- 1 **A.Charpentier** ,2010 : Statistique de l'assurance, HAL Id : cel-00550583, Université de Rennes 1 et Université de Montréal.
- 2 **J. Yeatman**,1998: Manuel international de l'assurance , éd Economica.
- 3 **Olga A. Vasechko, Michel Grun-Rehomme, Nouredine Benalagha**,(2009) :Modélisation de la fréquence des sinistres en assurance automobile, Bulletin français d'actuariat.
- 4 **P Therond, F Planchet, J Jacquemin**, 2010 :Modèles Financiers en assurance, Ed Economica, Paris.
- 5 **Landel .Jaunes etcharé-Serveau . Martine**,1997: Lexique juridique et pratique des termes d'assurance, édition l'Argus,
- 6 **Lambert Faivre. Y** ,1986: Droit des assurances, édition : Précis Dalloz,
- 7 **M. Denuit et A. Charpentier**,2004: Mathématique de l'Assurance Non-Vie, Tome I, Edition Economica.
- 8 **Mohamed Khordj**, quantification de l'incertitude liée au risque de provisionnement en assurance non-vie via les modèle stochastique, mémoire magister ,ENSSEA ,2015
- 9 **Meriem Riad**, "modèle de tarification optimal en assurance automobile dans le cadre d'un marché réglementé : cas de l'Algérie ", Revue d'économie et statistique appliquée (25) ,213-227.
- 10 **R .RIMI, A. LATRECHE, O. RIMI**, (2015). une évaluation empirique de la tarification de l'assurance Automobile en Algérie.RoaIktissadia Review.
- 11 **Rochelle .André** ,1999: Livre d'assurance : top pratique, paris édition.

- 12 **Stéphane, T.** (2012). *Data mining et statistique décisionnelle: l'intelligence des données*. Editions Technip.
- 13 **Sylvie .Chanh ; Jean.Péchinot.** ,2016: manuel des assurance automobile ,5eme éditions, l'agrus,paris
- 14 Assurance automobile, Le portail de l'assurance en Algérie, consulté le 17 avril 2019
- 15 Bulletin des Assurances n°21 .CNA.
- 16 Revue éditée n° 17 par CNA avril2017 « sinistralité automobile en Algérie vers de nouveaux critères de tarifications
- 17 <http://www.cna.dz>.
- 18 <http://www.protectioncivile.dz/?controller=listing&action=contenu&idr=%2043>
- 19 <http://www.protectioncivile.dz/>