CNAM 2002-2003 – 2léments de cours –
Bonus-malus et Crédibilité

Une situation fréquente en pratique est de disposer non pas d'un résultat mais de plusieurs.
Le cas se présente en assurance, par exemple :

- lors de la phase de tarification, la prime pure estimée, pour un même risque, peut varier en fonction de la finesse et des critères de segmentation.
- une fois le tarif établi, la comparaison des résultats obtenus avec les résultats attendus peut faire apparaître des écarts, font-ils partie de l'incertitude normale inhérente à l'activité, ou doivent-ils entraîner des correctifs tarifaires ?
- pour comparer les résultats obtenus avec les résultats attendus, encore faut-il apprécier correctement les résultats réels, qui dépendent pour une part de l'évaluation de règlements futurs (les provisions) et pour lesquels plusieurs méthodes, ou modèles, sont utilisables.

En pratique, l'objectif n'est pas de sélectionner un des résultats au détriment des autres, mais d'essayer de les combiner/réconcilier, en tenant compte du degré de confiance que l'on peut accorder à chacune de ces estimations.

Une application traditionnelle est appelée de manière générique « systèmes de bonus-malus ». Ils visent à pondérer l'appréciation du risque que l'on peut faire à priori (par exemple à l'issue d'une méthode de segmentation) avec l'appréciation du risque que l'on peut faire a posteriori sur la base de ses réalisations.

Les travaux sur les bonus-malus ont ensuite donné naissance à la théorie de la crédibilité, qui visent à modéliser les « meilleures » (au sens d'un modèle statistique sous-jacent) pondérations entre « vision à priori » et « vision à posteriori ».

Issus de la tarification, ces outils ont également été utilisés pour valider théoriquement des méthodes empiriques traditionnelles d'évaluation des provisions, et notamment la méthode Bornhuetter-Ferguson, seconde méthode la plus utilisée après celle des cadences de développement (chain-ladder).

1. Le système de bonus malus en assurance automobile.

La plupart des systèmes de bonus-malus techniques corrigent la prime pure d'un assuré au vu de l'historique de sa sinistralité.

Les caractéristiques générales des systèmes de bonus-malus en usage en assurance automobile sont les suivantes :

- une échelle de niveau de primes, la base 100 correspondant, dans chaque case du tarif, à la prime de référence ;
- la règle de bonification ou de pénalisation ne dépend pas de l'ancienneté de l'assuré ;
- cette règle ne prend en compte que le nombre de sinistres de la dernière année (ou des deux dernières années) ;
- il y a un niveau maximum de bonus et parfois un maximum de malus.

La clause-type française fixe un coefficient de tarification égal à 1,00 à l'origine, réduit de 5 % (ou 7 % pour les V.R.P.) par année sans sinistre, majoré de 25 % (ou 20 % pour les V.R.P.), avec une descente rapide à 1,00, après deux ans sans sinistres. Le niveau minimum est de 0,50 et le maximum de 3,50.

L'étude de l'évolution de la répartition des assurés entre les différents niveaux de primes relève de la théorie des chaînes de Markov.

Nous nous contenterons d'effectuer une étude simple afin d'apprécier le pouvoir discriminant de la clause de bonus malus.

Soit une population d'assurés comportant 80 % d'assurés ayant une fréquence d'accident de 5 % et 20 % d'assurés ayant une fréquence d'accident de 15 %, la survenance des accidents pour chacun des assurés étant un processus de Poisson.

Les sinistres sur cette population surviendront donc selon un processus de Poisson de paramètre $n \times (80\ \% \times 5\ \% + 20\ \% \times 15\ \% = 6\ \%)$ où n est l'effectif de la population. Leur origine, un bon ou un mauvais conducteur, est donc indiscernable.

Quelles sont les probabilité qu'un assuré, bon ou mauvais, n'ait pas de sinistre pendant un an et pendant deux ans ? Parmi ceux qui n'ont pas de sinistres, quelle proportion y a-t-il de bons ou mauvais conducteurs ?

		Un an	Proportion	Deux ans	Proportion
Aucun sinistre	Bon conducteur	95,1%	81,6%	90,5%	83,0%
	Mauvais conducteur	86,1%	18,4%	74,1%	17,0%
	Ensemble	93,3%	100,0%	87,2%	100,0%
Au moins un sinistre	Bon conducteur	4,9%	58,3%	9,5%	59,5%
	Mauvais conducteur	13,9%	41,7%	25,9%	40,5%
	Ensemble	6,7%	100,0%	12,8%	100,0%

Pour un assuré, ne pas avoir de sinistre en un an conduit à estimer sa fréquence à $(81{,}6\ \% \times 5{,}0\ \% + 18{,}4\ \% \times 15{,}0\ \%) = 6{,}84\ \%$, soit une diminution relative de 2 % par rapport à la fréquence a priori. La fréquence moyenne des autres conducteurs peut être estimée à 9,16 % (+ 31 %).

Pour un assuré, ne pas avoir de sinistre pendant deux ans conduit à estimer sa fréquence à $(83{,}0\ \% \times 5{,}0\ \% + 17{,}0\ \% \times 15{,}0\ \%) = 6{,}70\ \%$, soit une diminution relative de 4 %. La fréquence moyenne des autres conducteurs peut être estimée à 9,05 % (+ 29 %).

2. Etude de la prime modelée.

$\text{Gamma}_{\gamma,c}(\lambda) = \dfrac{c^{\gamma}}{\Gamma(\gamma)} \times \lambda^{\gamma-1} \times e^{-c\lambda}$
$\text{BN}_{\gamma,p}(k) = \dfrac{\Gamma(\gamma+k)}{\Gamma(\gamma)\times k!} \times p^{\gamma} \times (1-p)^{k}$

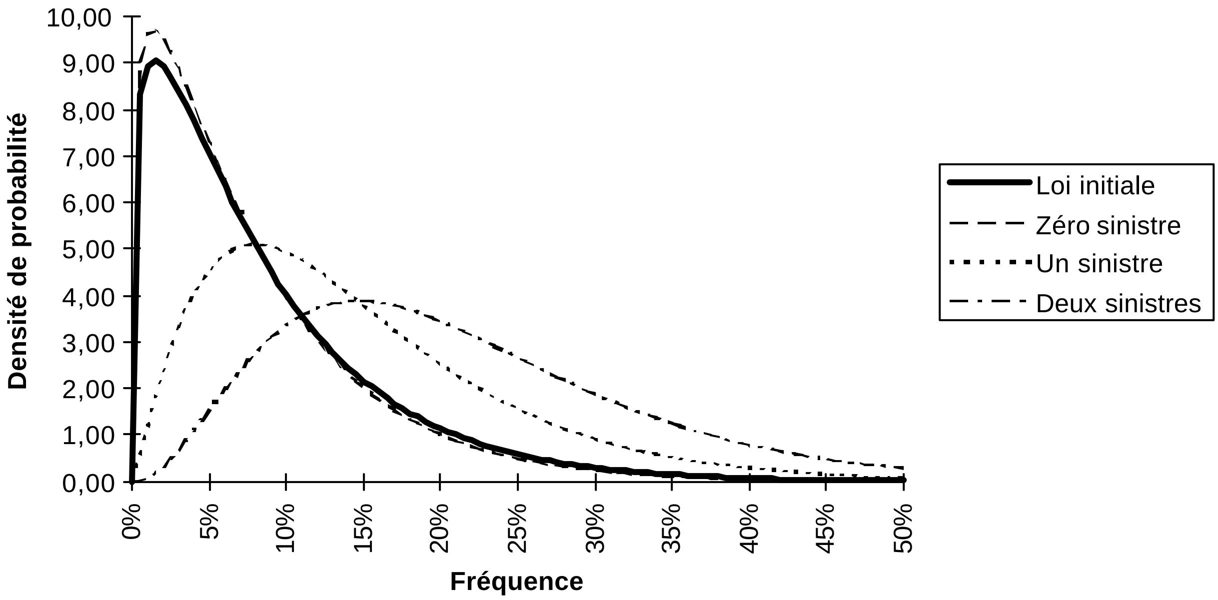
Lorsque le niveau du risque (le paramètre ***l*** de la loi de Poisson) est une réalisation d'une loi Gamma de paramètres ***g*** et *c*, la loi globale du nombre de sinistres survenant pendant une période *t* suit une loi binomiale négative de paramètres ***g*** et *p* = *c*/(*c*+*t*).

On démontre que si pendant une période *T* il survient *N* sinistres, alors le niveau du risque suit une loi Gamma de paramètres ***g*¢** = ***g*** + *N* et *c*¢ = *c*+ *T*.

Après *T* années d'observations où on a observé *N* sinistres, la prime pure devient : P = C × (***g*** + *N*)/(*c* + *T*), soit en divisant par *T* numérateur et dénominateur :

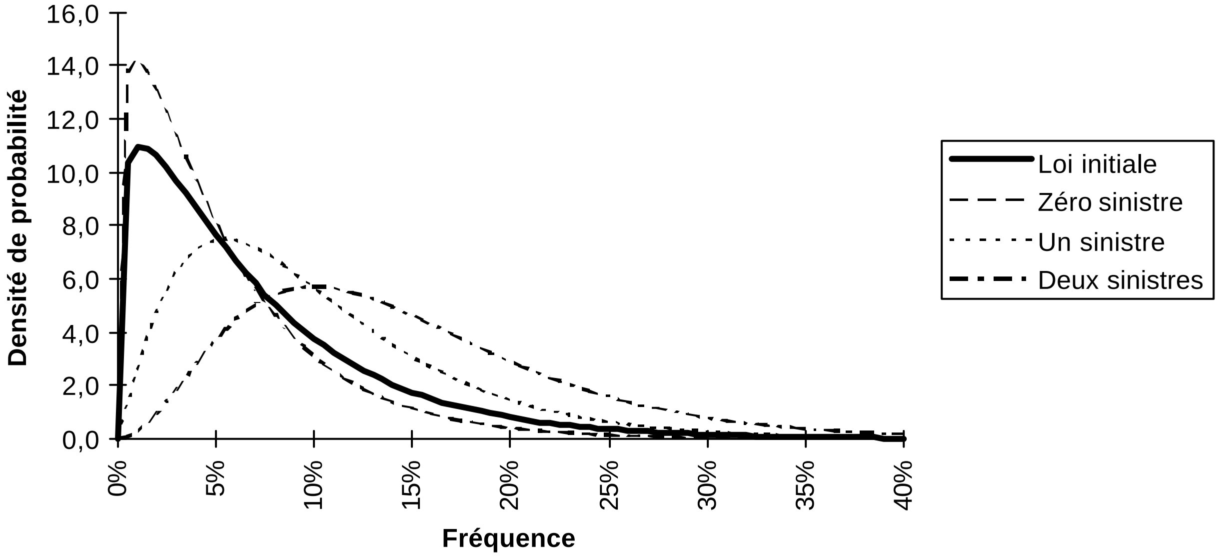
P = C × (*N*/*T* + ***g***/*T*)/(1 + *c*/*T*)

Ajustement de la distibution des fréquences suivant la sinistralité



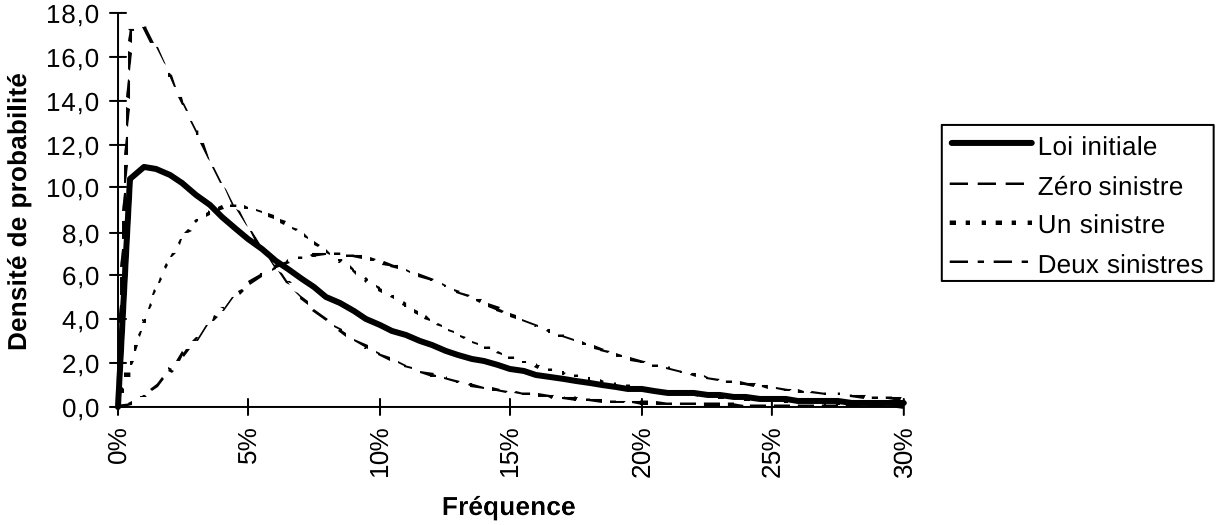
	proba	*c*	***g***	E(λ)	σ(λ)
Loi initiale		17	1,2	7,1 %	6,4 %
Zéro sinistre	93,4 %	18	1,2	6,7 %	6,1 %
Un sinistre	6,2%	18	2,2	12,2 %	8,2 %
Deux sinistres	0,4%	18	3,2	17,8 %	10,0 %

Ajustement de la distibution des fréquences suivant la sinistralité pendant cinq ans



	proba	*c*	*g*	E(λ)	σ(λ)
Loi initiale		17	1,2	7,1 %	6,4 %
Zéro sinistre	73,4 %	22	1,2	5,5 %	5,0 %
Un sinistre	20,0 %	22	2,2	10,0 %	6,7 %
Deux sinistres	5,0 %	22	3,2	14,5 %	8,1 %
Trois sinistres	1,2 %	22	4,2	19,1 %	9,3 %

Ajustement de la distibution des fréquences suivant la sinistralité pendant dix ans



	proba	c	g	E(λ)	σ(λ)
Loi initiale		17	1,2	7,1 %	6,4 %
Zéro sinistre	57,4 %	27	1,2	4,4 %	4,1 %
Un sinistre	25,5 %	27	2,2	8,1 %	5,5 %
Deux sinistres	10,4 %	27	3,2	11,9 %	6,6 %
Trois sinistres	4,1 %	27	4,2	15,6 %	7,6 %

3. La théorie de la crédibilité.

On s'intéresse à une variable aléatoire ${}_iX_j$, relative au risque i, observée en année j ; cela peut être, par exemple, la charge annuelle de sinistres, le nombre de sinistres ou le taux de sinistres à primes. Les caractéristiques de la variable aléatoire ${}_iX_j$ dépendent de la valeur d'un paramètre aléatoire θ_i, propre au risque i.

On cherche à évaluer ${}_iP_{n+1} = E({}_iX_{n+1})$, compte tenu des observations faites :
- d'une part sur le risque i, soit : ${}_iX_1, {}_iX_2, {}_iX_n$;
- d'autre part sur les risques de même nature.

L'estimateur de crédibilité est une fonction affine ou linéaire des observations faites sur le risque i.

Nous supposerons que les variables ${}_iX_j$ sont indépendantes et que, pour i fixé, elles sont identiquement distribuées. Nous avons alors deux variables aléatoires : $m(\theta_i) = E({}_iX_j)$ et $\sigma^2(\theta_i) = Var({}_iX_j)$.

Par exemple, lorsque les variables ${}_iX_j$ suivent une loi de poisson de paramètre λ_i, $E({}_iX_j) = Var({}_iX_j) = \lambda_i$.

En général, on ne connaît pas la fonction de structure mais on suppose qu'existent :
$m = E(m(\theta_i))$
$a = Var(m(\theta_i))$
$s^2 = E(\sigma^2(\theta_i))$

Par exemple, lorsque les variables ${}_iX_j$ suivent une loi de poisson de paramètre λ_i, et que la variable aléatoire λ_i suit une loi Gamma de paramètres g et c, $m = g/c$, $a = g/c^2$ et $s^2 = g/c$.

Nous avons alors les propositions suivantes (dues à Bühlmann) que nous livrons sans démonstration :

> La meilleure approximation de $m(\theta_i)$ au sens des moindres carrés, par une fonction affine des observations $({}_iX_1, {}_iX_2, {}_iX_n)$ est la variable aléatoire :

> $(1\text{-}Z_n) \times m + Z_n \times ({}_iX_1 + {}_iX_2 + \ldots + {}_iX_n)$ où $Z_n = a/(a+s^2/n)$
> Z_n est appelé le facteur de crédibilité.
> L'erreur d'approximation est $\leq (1\text{-}Z_n) \times a$

Lorsque les variables ${}_iX_j$ suivent une loi de poisson de paramètre λ_i, et que la variable aléatoire λ_i suit une loi Gamma de paramètres g et c, m = g/c et $Z_n = 1/(1+c/n)$. On retrouve la même formule que précédemment.

De façon analogue, on montre la proposition suivante, en se plaçant dans la classe des fonctions linéaires.

> La meilleure approximation au sens des moindres carrés de $m(q_i)$ par une fonction linéaire des observations est la variable aléatoire :
> $(a + m^2)/(s^2 + (a + m^2) \times n) \times ({}_iX_1 + {}_iX_2 + \ldots + {}_iX_n)$
> L'erreur d'approximation est $\leq (s^2 \times (a + m^2))/(s^2 + (a + m^2) \times n)$

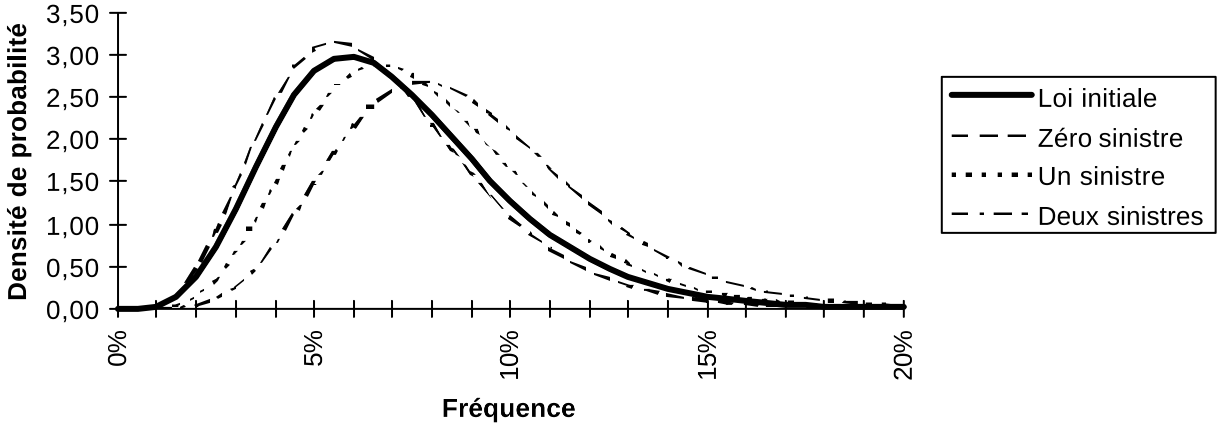
On voit que l'erreur est supérieure à celle de la proposition précédente. Par ailleurs l'espérance de cette loi est $(a + m^2))/(s^2 + (a + m^2) \times n) \times n \times m < E(m(q_i)) = m$.

Si $m(q_i)$ est la prime pure d'un contrat, la prime calculée par cette approximation est en moyenne inférieure à la vraie prime pure, ce qui évidemment n'est pas souhaitable.

4. Etude des flottes.

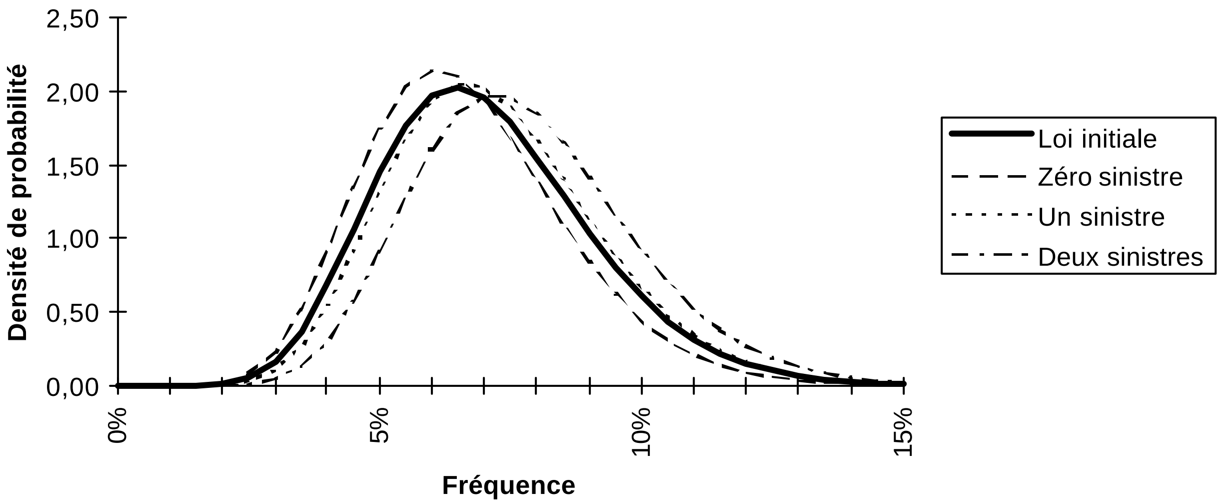
Soit une flotte de n véhicules, la survenance des sinistres suivant pour chacun d'entre eux un processus de Poisson de paramètre l, l étant pour chacun une réalisation d'une loi Gamma de paramètres g et c propres à la flotte. On démontre que la survenance des sinistres au sein de la flotte suit alors un processus de Poisson de paramètre $l¢$ propre à la flotte, $l¢$ étant la réalisation d'une loi Gamma de paramètres g et nc. Le nombre total de sinistres survenant pendant une période t suit alors une loi binomiale négative de paramètres g et $p = nc/(nc+t)$.

Ajustement de la distibution des fréquences suivant la sinistralité
Flotte de cinq véhicules



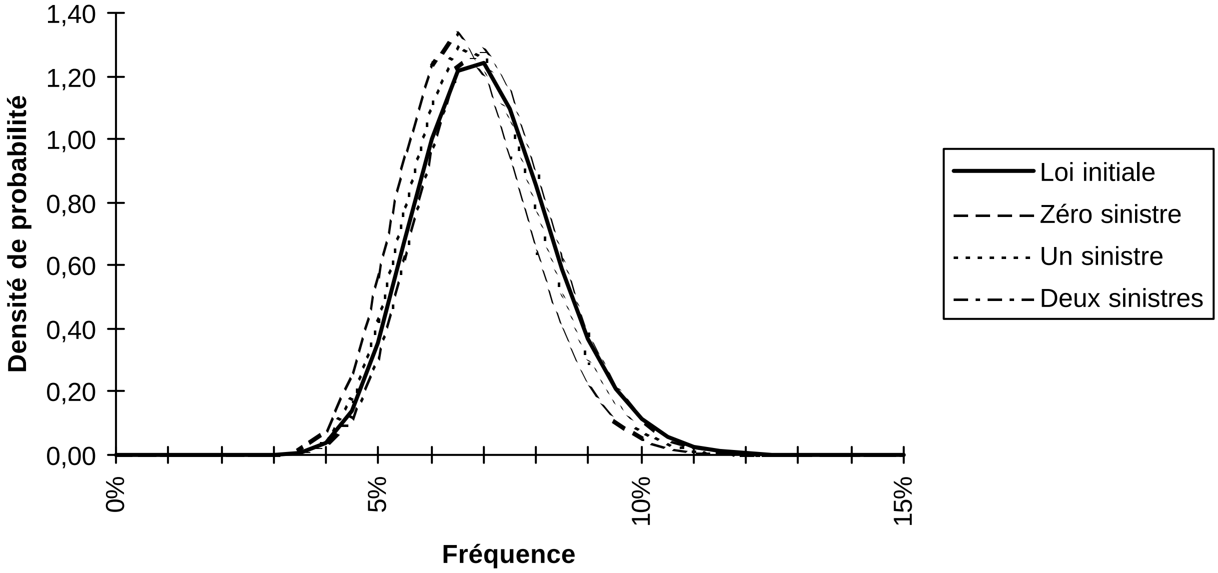
	proba	E(λ)	Ecart
Loi initiale		7,1%	
Zéro sinistre	71,0%	6,7%	-5,6%
Un sinistre	23,7%	7,8%	10,2%
Deux sinistres	4,6%	8,9%	25,9%
Trois sinistres	0,7%	10,0%	41,7%

Ajustement de la distibution des fréquences suivant la sinistralité
Flotte de dix véhicules



	proba	E(λ)	Ecart
Loi initiale		7,1%	
Zéro sinistre	50,4%	6,7%	-5,6%
Un sinistre	33,6%	7,2%	2,3%
Deux sinistres	12,1%	7,8%	10,2%
Trois sinistres	3,1%	8,3%	18,1%
Quatre sinistres	0,7%	8,9%	25,9%

Ajustement de la distibution des fréquences suivant la sinistralité
Flotte de 25 véhicules



	proba	E(λ)	Ecart
Loi initiale		7,1%	
Zéro sinistre	18,0 %	6,7 %	-5,6 %
Un sinistre	30,0 %	6,9 %	-2, 4%
Deux sinistres	25,8 %	7,1 %	0,7 %
Trois sinistres	15,3 %	7,3 %	3,9 %
Quatre sinistres	7,0 %	7,6 %	7,0 %
Cinq sinistres	2,7 %	7,8 %	10,2 %
Six sinistres	0,9 %	8,0 %	13,3 %
Sept sinistres	0,2 %	8,2 %	16,5 %