Mathématiques actuarielles IARD-1
ACT-2005
Notes de cours

Gabriel Crépeault-Cauchon
Nicholas Langevin

29 septembre 2018

Table des matières

Résumé

Ce document résume les notes de cours prises en classe dans le cours de Mathématiques actuarielles IARD-1, ainsi que des notions prises du livre *LOSS MODELS - From Data to Decisions, 4th edition*.

Chapitre 1

Rappel sur les notions de probabilité et statistiques

1.1 Quantités à savoir

Raw moments On représente le k^{e} moment par μ'_k, soit

$$\mu'_k = E\left[X^k\right] \tag{1.1}$$

Moments centraux Le k^{e} moment central est représenté par

$$\mu_k = E\left[(X-\mu)^k\right] \tag{1.2}$$

Exemple 1.1.1 Quelques exemples de moments centraux

La variance est le 2^{e} moment central :

$$Var(X) = \mu_2 = E\left[(X-\mu)^2\right]$$

Le 3^{e} moment centré, qui est utilisé pour calculer le coefficient d'asymétrie :

$$\mu_3 = E\left[(X-\mu)^3\right]$$

Coefficient d'asymétrie Le coefficient d'asymétrie, aussi appelé *skewness*, est représentée par

$$S_k = \frac{\mu_3}{\sigma^2} \tag{1.3}$$

Soit le 3ᵉ moment standarisé. Si $S_k = 0$, alors la distribution tend vers une loi normale.

Coefficient d'applatissement Le coefficient d'applatissement, aussi appelé *Kurtosis*, se définit par

$$\text{Kurtosis} = \frac{\mu_4}{\sigma^4} \tag{1.4}$$

Cette quantité permet de mesurer l'épaisseur de l'aile (*tail*) de la distribution. Si $E\left[z^4\right] = 3$, alors la distribution tend vers une loi normale $N(\mu, \sigma^2)$.

1.2 La loi normale

La fonction génératrice des moments

$$\begin{aligned} M_x(t) &= M_x(0) + \frac{M_x' t}{1!} + \frac{M_x'' t^2}{2!} + \ldots + \frac{M_x^{(n)t^n}}{n!} \\ &= 1 + \frac{E[x]t}{1!} + \frac{E[x^2]t^2}{2!} + \ldots + \frac{E[x^n]t^n}{n!} \end{aligned}$$

On pose : $c_k = \frac{E[x^n]}{n!}$ alors,

$$E[x^k] = C_k k! \tag{1.5}$$

1.3 Queue de distribution

1. Sois $f_1(x)$ une fonction tels que les 3 premiers moment existe : $E[x^4] = \infty$
2. Sois $f_2(x)$ une fonction tels que les 2 premiers moment existe : $E[x^2] = \infty$

Alors,

$$\lim_{x \to \infty} r(x) = \frac{f_1(x)}{f_2(x)} = \begin{cases} \infty, \ f_1(x) \text{a une aile plus lourde que} f_2(x) \\ 0, \ f_2(x) \text{a une aile plus lourde que} f_1(x) \end{cases}$$

Exemple 1.3.1

Sois $f_{x_1}(x_1) \sim pareto(\alpha, \theta)$ et $f_{x_2}(x_2) \sim gamma(\alpha, \lambda)$

$$\begin{aligned}
\lim_{x\to\infty} r(x) &= \lim_{x\to\infty} \frac{f_{x_1}(x_1)}{f_{x_2}(x_2)} \\
&= \frac{\frac{\alpha\theta^{\alpha}}{(x+\theta)^{\alpha+1}}}{\lambda^{\alpha} x^{\alpha-1} e^{-\lambda x}} \\
&= C \frac{e^{-\lambda x}}{x^{\alpha-1}(x+\theta)^{\alpha+1}} \\
&= \infty
\end{aligned}$$

La pareto a une queue plus lourde que la gamma

La fonction de hasard

$$h(x) = \frac{f(x)}{s(x)} \tag{1.6}$$

Si à partir de M, $h(x)$ est décroissante $\Leftrightarrow$ $f(x)$ décroit trop lentement, alors $f(x)$ à une aile lourde.

Chapitre 2

Types de contrats et primes

Dans cette section [1], on définit deux nouvelles variables

Définition 2.0.1 *Per-Loss* Variable

Soit X le montant des dommages d'une réclamation. On peut définir Y^L comme le montant payé par l'assureur lors de toute perte. Mathématiquement,

$$Y^L = g(X) \quad , Y^L \sim f_X(x) \tag{2.1}$$

Définition 2.0.2 *Per-Payment* variable

Cette variable se définit plutôt comme le montant qui sera payé par l'assureur (i.e le montant de la perte, sachant qu'il y aura un paiement). Alors,

$$Y^P = g(X) \quad , Y^P \sim \frac{f_X(x)}{S_X(x)} \tag{2.2}$$

Y^P n'a donc pas de probabilité définie à $y = 0$ (puisque ce n'est pas possible, sachant que x est assez grand pour qu'il y ait un paiement.

1. Section 8.1 à 8.5 dans le livre

2.1 Contrat avec limite

On analyse la fonction de perte d'un contrat avec une limite supérieure u. Alors, on définit Y comme

$$Y = (X \wedge u) = \min(X, u) = \begin{cases} x & , x \leq u \\ u & , x > u \end{cases} \tag{2.3}$$

Aussi,

$$\begin{aligned} E[Y] &= E[\min(X, u)] \\ &= \int_0^u x f_X(x)dx + u \int_u^\infty f_X(x)dx \\ &= \int_0^u x f_X(x)dx + uS(u) \end{aligned}$$

2.1.1 Cas avec inflation

On a vu le contrat avec perte limitée, mais sans parler d'inflation. Supposons qu'on a le scénario <u>avec inflation</u>, où $X' = (1+r)X$ et r représente le taux d'inflation appliqué au montant de perte. Alors, il faut ajuster u pour tenir compte de l'inflation.

$$\begin{aligned} Y &= \begin{cases} (1+r)X & , (1+r)X \leq u \\ u & , (1+r)X > u \end{cases} \\ &= \begin{cases} (1+r)X & , x \leq \frac{u}{1+r} \\ u & , x > \frac{u}{1+r} \end{cases} \end{aligned}$$

Définition 2.1.1 Prime Limited Loss sous l'inflation

$$E\left[X' \wedge u\right] = (1+r)E\left[X \wedge \frac{u}{1+r}\right] \tag{2.4}$$

Démonstration.

$$
\begin{aligned}
E[Y] &= \int_0^{\frac{u}{1+r}} (1+r)xf_X(x)dx + u\int_{\frac{u}{1+r}}^{\infty} f_X(x)dx \\
&= (1+r)\int_0^{\frac{d}{1+r}} f f_X(x)dx + sS_X\left(\frac{u}{1+r}\right) \\
&= (1+r)\left(\int_0^{\frac{u}{1+r}} xf_X(x)dx + \frac{u}{1+r}S_X\left(\frac{u}{1+r}\right)\right) \\
&= (1+r)E\left[X \wedge \frac{u}{1+r}\right]
\end{aligned}
$$

□

2.2 Contrat avec déductible ordinaire

Soit un contrat avec perte X, où on paie 0 si $X \leq d$ et $X - d$ si $X > d$, où d est le déductible ordinaire. Alors,

$$
Y = (X-d)_+ = \max(X-d, 0) = \begin{cases} 0 & ,X \leq d \\ X-d & ,X > d \end{cases} \tag{2.5}
$$

2.2.1 Remarques

(1) X est une variable aléatoire continue

(2) mais Y^L est une variable aléatoire mixte, car

$$
\Pr\left(Y^L = y\right) = \begin{cases} F_X(d) & ,y = 0 \\ f_X(y+d) & ,y > 0 \end{cases}
$$

(3) La fonction de répartition de Y^L est définie par

$$
F_Y(y) = \begin{cases} F_X(d) & ,y = 0 \\ F_X(d) + \int_0^y f_X(u+d)du & \\ \cancel{F_X(d)} + F_X(y+d) \cancel{- F_X(d)} & \\ F_X(y+d) & y > 0 \end{cases}
$$

(4) les fonctions de survie $S(y)$ et fonction de hasard $h(y)$ sont définies par

$$
\begin{aligned}
S_Y(y) &= 1 - F_Y(y) \\
h(y) &= \frac{f_Y(y)}{S_Y(y)}
\end{aligned}
$$

(5) La prime de ce contrat est appelée la prime *stop-loss*, qui représente l'espérance de ce contrat.

2.2.2 Prime Stop-Loss

On veut calculer l'espérance du contrat avec déductible ordinaire :

$$
\begin{aligned}
E[Y] &= E\left[(X-d)_+\right] \\
&= \int_d^\infty (x-d)f(x)dx \\
&= \underbrace{\int_d^\infty xf(x)dx}_{\text{Intégration par partie}} \; - d\int_d^\infty f(x)dx \\
&= -xS(x)\Big|_d^\infty + \int_d^\infty S(x)dx - sS(d) \\
&= 0 \cancel{+S(d)} + \int_d^\infty S(x)dx \cancel{-S(d)} \\
&= \int_d^\infty S(x)dx
\end{aligned}
$$

Définition 2.2.1 Prime Stop-Loss

On peut définir la prime Stop-Loss comme

$$E\left[(X-d)_+\right] = E\left[X\right] - E\left[X \wedge d\right] \tag{2.6}$$

Démonstration. À compléter □

2.2.3 Cas avec inflation

$$Y = \begin{cases} 0 & ,X' \leq d \\ X' - d & ,d' > d \end{cases}$$

où $X' = (1+r)X$. On peut travailler seulement avec le X initial :

$$Y = \begin{cases} 0 & ,x \leq \frac{d}{1+r} \\ (1+r)X - d & ,x > \frac{d}{1+r} \end{cases}$$

Et la prime *Stop-Loss*,

$$\begin{aligned} E[Y] &= \int_{\frac{d}{1+r}}^{\infty} ((1+r)x - d) f_X(x) dx \\ &= (1+r) \int_{\frac{d}{1+r}}^{\infty} x f_X(x) dx - \frac{d}{1+r} \int_{\frac{d}{1+r}}^{\infty} f_X(x) dx \end{aligned}$$

Définition 2.2.2 Prime Stop-Loss sous le scénario d'inflation

$$E\left[(X'-d)_+\right] = (1+r)\left(E[X] - E\left[X \wedge \frac{d}{1+r}\right]\right) \tag{2.7}$$

Démonstration.

$E[Y] = (1+r) \int_{\frac{d}{1+r}}^{\infty} x f_X(x) dx - d \int_{\frac{d}{1+r}}^{\infty} f_X(x) dx$

En ajoutant un terme,

$$= (1+r) \underbrace{\int_0^{\frac{d}{1+r}} x f_X(x) dx + \int_{\frac{d}{1+r}}^{\infty} x f_X(x) dx}_{E[X]} - \underbrace{\int_0^{\frac{d}{1+r}} x f_X(x) dx - \frac{d}{1+r} \int_{\frac{d}{1+r}}^{\infty} f_X(x) dx}_{E\left[X \wedge \frac{d}{1+r}\right]}$$

$$= E[X] - E\left[X \wedge \frac{d}{1+r}\right]$$

$\square$

2.3 Mean Excess Loss

Le Mean Excess Loss représente la perte excédentaire à d, sachant que $X > d$. Mathématiquement,

$$Y = \begin{cases} 0 & ,x \leq d \\ X - d & ,x > d \end{cases}$$

où $X \sim \frac{f(x)}{S(d)}, x \geq d$.

2.3.1 Remarques

(1) La fonction de densité de Y est représentée par

$$\Pr\left(Y^P = y\right) = \left\{ \frac{f_X(y+d)}{S(d)} \quad ,y > 0 \right.$$

(2) Y est une v.a. continue

(3) les fonctions de survie et de hasard sont, respectivement :

$$\begin{aligned} h(y) &= \frac{f_Y(y_d)}{S_Y(y+d)} \\ S(y) &= \frac{S_X(y+d)}{S(d)} \end{aligned}$$

(4) l'espérance de la v.a. Y^P[2] (mean excess loss) est définie par

$$\begin{aligned} e(d) &= E\left[Y^P\right] \\ &= E\left[X - d | X > d\right] \\ &= \int_d^\infty (x-d)\frac{f(x)}{S(d)}dx \end{aligned}$$

On note donc

$$e(d)S(d) = \underbrace{\int_d^\infty (x-d)f(x)dx}_{\text{Prime Stop-Loss}} \tag{2.8}$$

2. Le P en exposant signifie *per pay*, qui indique que Y est conditionnée sur X.

Définition 2.3.1 Loss Eliminating Ratio

Le Loss Eliminating Ratio (LER), nous permet d'obtenir le pourcentage de perte qu'on ne paiera pas grâce au déductible d :

$$ELR = \frac{E[X] - E[(X-d)_+]}{E[X]}$$

Mais on sait que

$$E[(X-d)_+] = E[X] - E[X \wedge d]$$

Alors,

$$ELR = \frac{E[X \wedge d]}{E[X]} \tag{2.9}$$

2.4 Contrat avec Déductible franchise

Soit Y le contrat qui a un déductible franchise. Dans ce type de contrat, on va payer l'intégralité de la perte, lorsque celle-ci dépassera un déductible. Mathématiquement,

$$Y = \begin{cases} 0 & ,X \leq d \\ X & ,X > d \end{cases}$$

Caractéristiques

(1) La fonction de densité est définie par

$$f_Y(y) = \begin{cases} F_X(d) & ,y = 0 \\ f_X(y) & ,y > 0 \end{cases}$$

(2) La fonction de répartition de ce contrat est

$$F_Y(y) = \begin{cases} F_X(d) & ,y = 0 \\ F_X(d) & ,0 < y \leq d \\ \frac{F_X(y) - F_X(d)}{S_X(d)} & ,y > d \end{cases}$$

2.5 Contrat avec déductible ordinaire et limite sous l'inflation

On peut combiner les contrats vu plus tôt, comme c'est le cas ici : ce contrat prévoit un déductible ordinaire d ainsi qu'une limite au contrat u. De plus, on s'intéresse au scénario sous l'inflation.Alors,

$$Y = \begin{cases} 0 & ,X \leq \frac{d}{1+r} \\ (1+r)X - d & ,\frac{d}{1+r} \leq X \leq \frac{u}{1+r} \\ u = d & ,X > \frac{u}{1+r} \end{cases}$$

Et l'espérance est

Définition 2.5.1 Prime d'un contrat avec déductible d et limite u

$$E[Y] = (1+r)\left(E\left[X \wedge \frac{u}{1+r}\right] - E\left[X \wedge \frac{d}{1+r}\right]\right) \tag{2.10}$$

Démonstration. On va utiliser $Y = Y_1 - Y_2$ pour prouver (2.10). On définit Y_1 et Y_2 :

$$Y_1 = \begin{cases} 0 & ,X \leq \frac{d}{1+r} \\ (1+r)X - d & ,X > \frac{d}{1+r} \end{cases}$$
$$Y_2 = \begin{cases} 0 & ,X \leq \frac{u}{1+r} \\ (1+r)X - u & ,X > \frac{u}{1+r} \end{cases}$$

On connait leur espérance respective,

$$E[Y_1] = (1+r)\left(E[X] - E\left[X \wedge \frac{d}{1+r}\right]\right)$$
$$E[Y_2] = (1+r)\left(E[X] - E\left[X \wedge \frac{u}{1+r}\right]\right)$$

Alors,

$$\begin{aligned} E[Y] &= E[Y_1] - E[Y_2] \\ &= (1+r)\left(E\left[X \wedge \frac{u}{1+r}\right] - E\left[X \wedge \frac{d}{1+r}\right]\right) \end{aligned}$$

$\square$

$$E[Y] = \int_{\frac{d}{1+r}}^{\frac{u}{1+r}} ((1+r)x-d) f_X(x)dx + \int_{\frac{u}{1+r}}^{\infty} (u-d) f_X(x)dx$$
$$= (1+r)\left(E\left[X \wedge \frac{u}{1+r}\right] - E\left[X \wedge \frac{d}{1+r}\right]\right)$$

2.6 Co-assurance

Soit α le pourcentage de co-assurance.

$$Y = \begin{cases} 0 & ,X \leq d \\ X-d & ,X > d \end{cases}$$

$$Y^C = \begin{cases} 0 & ,X \leq d \\ \alpha(X-d) & ,X > d \end{cases}$$

Alors,

$$E\left[Y^C\right] = \alpha E[Y] \tag{2.11}$$

2.7 Cas général

Voici une formule générale, avec un déductible d, un taux d'inflation r, un pourcentage de coassurance de α et une limite de u :

$$Y^L = \begin{cases} 0 & , \ x \leq \frac{d}{1+r} \\ \alpha\Big((1+r)x-d\Big) & , \ \frac{d}{1+r} \leq x \leq \frac{u}{1+r} \\ \alpha(u-d) & , \ x > \frac{u}{1+r} \end{cases} \tag{2.12}$$

Chapitre 3

Estimation non-paramétrique

À savoir pour examen partiel

- ✔ Loi normale
- ✔ Loi gamma
- ✔ Loi poisson
- ✔ Loi binomiale
- ✔ intégration par partie poru la loi exponentielle
- ✔ fonction densité, moyenne, variance et fgm

3.1 Terminologie

Pour $X_1, \ldots, X_n$ qui sont *iid*.
On a

$$\begin{aligned}
E\left[g(X)\right] &= \int g(x)dF(x) \\
&= \int g(x)f_X(x)dx \\
&= \int g(x)F_n(x) \\
&= \frac{1}{n}\sum_{i=1}^{n} g(x_i)
\end{aligned}$$

... est un estimateur non-paramétrique

3.2 Estimation de données complètes

On cherche à estimer $F(t)$ ou $S(t)$, lorsque nos données sont complètes (i.e. $x_1,...,x_n$ qui sont *iid*). Alors, l'estimateur non paramétrique pour $F(t)$:

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} 1_{[x_i \leq t]} \tag{3.1}$$

où $1_{[\cdot]}$ représente une fonction indicatrice.
$F_x(t)$ aura donc la forme suivante :

$$F_n(t) = \begin{cases} 0 & ,t < x_{(1)} \\ \frac{1}{n} & ,x_{(1)} \leq t < x_{(2)} \\ \frac{2}{n} & ,x_{(2)} \leq t < x_{(3)} \\ \ldots & \\ 1 & ,t \geq x_{(n)} \end{cases} \tag{3.2}$$

où $t \in [0, x_{(n)}]$.

Remarques

(1) Lorsque $F_n(t) \to F(t)$, alors

(2) Puisqu'on a

$$\sum_{i=1}^{n} 1_{[x_i \leq t]} \sim Bin(n, \Pr(X \leq t))$$

Alors,

$$\begin{aligned} E\left[F_n(t)\right] &= \frac{1}{n} nF(t) \\ &= F(t) \quad \text{(C'est un estimateur sans biais)} \\ Var\left(F_n(t)\right) &= \frac{1}{n^2} nF(t)S(t) \\ &= \frac{F(t)S(t)}{n} \\ &\underset{n\to\infty}{=} 0 \end{aligned}$$

3.2.1 Estimation de donnée groupées (Fonction OGIVE)

CETTE MATIÈRE NE SERA PAS TESTÉE À L'EXAMEN

Dans certains contextes, on a n données qui sont groupées en intervalle. La fonction OGIVE permet d'interpoler entre 2 points x_i et $_{i+1}$.

$$c_{j-1} \leq x \leq c_j$$
$$F_n(c_{j-1}) \leq F_n(x) \leq F_n(c_j)$$

La formule est

$$F_n^{\text{OGIVE}}(x) = \frac{c_j - x}{c_j - c_{j-1}} F_n(c_j - 1) + \frac{x - c_{j-1}}{c_j - c_{j-1}} F_n(c_j) \tag{3.3}$$

Remarques

(1) Si $x = c_{j-1}$,

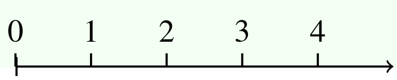
$$F_n(c_{j-1}) = F_n^{\text{OGIVE}}(c_{j-1})$$

(2) Si $x = c_j$,

$$F_n(c_j) = F_n^{\text{OGIVE}}(c_j)$$

Exemple 3.2.1 Exemple concret

L'assureur a groupé $n = 10$ données en intervalles.



Alors,

$$
\begin{aligned}
F_n(t) &= \frac{1}{n}\sum_{i=1}^{n} I[x_i \leq t] \\
F_n(1) &= \frac{2}{10} \\
F_n(2) &= \frac{5}{10} \\
F_n(3) &= \frac{9}{10} \\
F_n(4) &= \frac{10}{10} \\
&= 1
\end{aligned}
$$

En dérivant (3.3), on obtient

$$
\begin{aligned}
f_n(x) &= \frac{\partial}{\partial x} F_x(x) \\
&= \frac{1}{c_j - c_{j-1}} F_x(c_j) - \frac{1}{c_j - c_{j-1}} F_n(c_{j-1}) \\
&= \frac{F_n(c_j) - F_n(c_{j-1})}{c_j - c_{j-1}}
\end{aligned}
\tag{3.4}
$$

3.3 Estimation de la fonction de survie

Soit $S_n(t)$ la fonction de survie empirique. Alors,

$$\begin{aligned} S_n(t) &= 1 - F_n(t) \\ &= \frac{n}{n} - \frac{1}{n}\sum_{i=1}^{n} 1_{\{x_i \leq t\}} \\ &= \frac{1}{n}\sum_{i=1}^{n}\left(1 - 1_{\{x_i \leq t\}}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} 1_{\{x_i > t\}} \end{aligned}$$

3.4 Estimateur Kaplan-Meier

3.5 l'approche conditionnelle pour estimer $S(t)$

$$\begin{aligned} S(t) &= \frac{S(t_1)}{S(t_0)} \times \frac{S(t_2)}{S(t_1)} \times \ldots \times \frac{S(t)}{S(t_{i-1}} \\ &= \frac{S(t)}{\underbrace{S(t_0)}_{1}} \\ &= S(t) \end{aligned}$$

Alors,

$$S(t) = \prod_{j \leq t} p_j \tag{3.5}$$

Ça nous suggère un autre estimateur pour $S(t)$:

$$\hat{S}(t) = \prod_{j \leq t}(1 - \hat{q}_j) \tag{3.6}$$

Ceci est l'approche conditionnelle pour estimer $S(t)$. Et si jamais on a des données complètes, $S_n(t) = \hat{S}(t)$.