Mathématiques actuarielles IARD-1
ACT-2005
Notes de cours

Gabriel Crépeault-Cauchon
Nicholas Langevin

6 octobre 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 3

Quantités des distributions à connaître

3.1 Moments

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

$$\mu'_k = E\left[X^k\right] \tag{3.1}$$

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

$$\mu_k = E\left[(X-\mu)^k\right] \tag{3.2}$$

Exemple 3.1.1 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 :

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

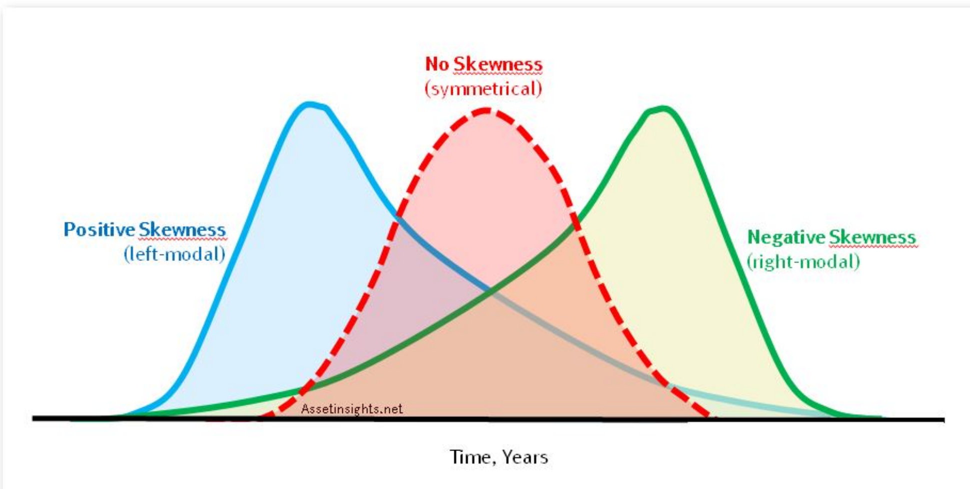
Coefficient de variation

$$CV = \frac{\sigma}{E[X]} \tag{3.3}$$

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

$$\gamma_1 = \frac{\mu_3}{\sigma^3} \tag{3.4}$$

Soit le 3e moment standarisé. Si $S_k = 0$, alors la distribution tend vers une loi normale, telle qu'on le voit sur la figure ci-dessous :



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

$$\text{Kurtosis} = \frac{\mu_4}{\sigma^4} \tag{3.5}$$

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 $N(\mu, \sigma^2)$.

Mean Excess Loss On définit la variable aléatoire Y^P, qui représente le montant de perte en excès d'un déductible d, sachant que la perte est au delà de ce montant. On peut définir l'espérance des coûts de cette variable alétaoire :

$$e(d) = E\left[Y^P\right] = E\left[X - d | X \geq d\right] = \frac{\int_d^\infty S_X(x)}{S_X(d)} \tag{3.6}$$

Note : cette variable est dite tronquée à gauche et *shifted*. On entend parfois aussi *Per-payment*

Left censored and shifted variable Soit la v.a. Y^L, qui représente le montant payé par l'assureur *par perte*. La variable est donc dite *censurée à gauche et shifted*. On peut aussi en calculer l'espérance :

$$E\left[Y^L\right] = E\left[(X-d)_+\right] = \int_d^{\infty} (x-d) f_X(x) dx \tag{3.7}$$

De plus, on peut facilement déduire la relation suivante :

$$E\left[(X-d)_+\right] = e(d)(1-F_X(d))$$

Limited Loss Variable Finalement, on peut définir la variable Y, qui représente le paiement de l'assureur avec une limite de u à la police. Son espérance est définie par

$$E\left[X \wedge u\right] = \int_0^u f_X(x) dx \tag{3.8}$$

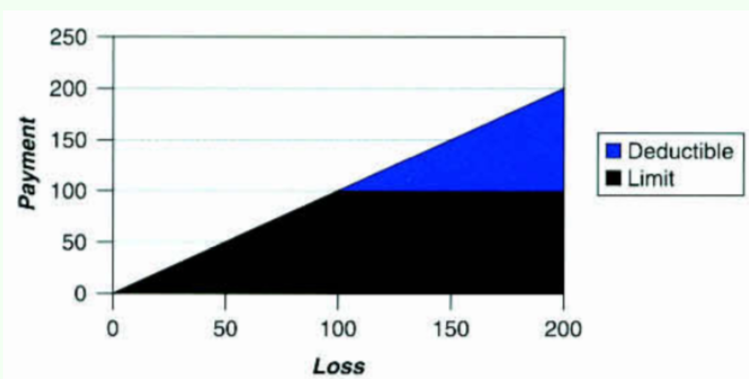
À l'aide d'intégration par partie, on peut trouver la forme suivante :

$$E\left[X \wedge u\right] = \int_0^u S_X(x) dx$$

La v.a. Y est dite *censurée à droite et shifted*

Exemple 3.1.2 lien entre le déductible et la limite

On peut faire le lien entre le déductible et la limite :



3.4 Queue de distribution

3.4.1 Classification selon les moments

- On peut déterminer si une distribution a une *heavy-tail* en vérifiant si ses moments existent.
- On peut aussi comparer des distributions entre-eux en utilisant des quantités standarisées, telles que le Coefficient de variation, le coefficient d'asymétrie (*skewness*) ou encore le coefficient d'applatissement (*Kurtosis*)

3.4.2 Classification selon les comportements limites des ailes de distribution

- On peut faire le ratio de deux distributions avec leurs fonction de survie ($S(x)$) ou leur fonction f de densité pour vérifier laquelle des 2 a la plus grosse aile de distribution (*tail*).

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 3.4.1

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

3.4.3 Classification basée sur la fonction de Hazard

Définition 3.4.1 *Hazard rate function*

La fonction de hazard (aussi appelée force de mortalité ($\mu(x)$) ou le failure rate ($\lambda(x)$), est définie par

$$h_X(x) = \frac{f_X(x)}{S_X(x)} \tag{3.9}$$

On peut aussi exprimer la fonction $h_X(x)$ comme

$$h_X(x) = -\ln(S_X(x))$$

Soit une distribution ayant fonction de densité $f_X(x)$ et fonction de hazard $h_X(x)$. Alors,

- Si $h(x) \nearrow$, *light-tailed*
- Si $h(x) \searrow$, *heavy-tailed*
- Note : on peut aussi comparer les distributions entre-elles : si une distribution voit son $h_1(x)$ augmenter plus rapidement que l'autre (i.e. $h_2(x)$), alors la deuxième distribution a une aile de distribution plus lourde.

Chapitre 8

Fréquence et sévérité avec modifications aux contrats

8.2 Déductibles

8.2.1 Ordinary deductible

Définition

Soit un contrat d'assurance avec déductible d. Lors d'une perte, l'assureur va payer tout montant en excédent du montant d. Alors, pour la variable *per-payment*,

$$Y^L = (X-d)_+ = \begin{cases} 0 & ,X \leq d \\ X-d & ,X > d \end{cases} \tag{8.1}$$

$$Y^P = (X-d)_+ = \begin{cases} \text{Non-défini} & ,X \leq d \\ X-d & ,X > d \end{cases} \tag{8.2}$$

Fonctions reliées

Et on peut aussi déduire toutes les fonctions qui y sont reliées :

$$f_{Y^P}(y) = \frac{f_X(y+d)}{S_X(d)}$$

$$S_{Y^P}(y) = \frac{S_X(y+d)}{S_X(d)}$$

$$F_{Y^P}(y) = \frac{F_X(y+d) - F_X(d)}{S_X(d)}$$

$$h_{Y^P}(y) = h_X(y+d)$$

Pour la v.a. Y^L *per-loss* [1],

$$f_{Y^L}(y) = f_X(y+d)$$

$$S_{Y^L}(y) = S_X(y+d)$$

$$F_{Y^L}(y) = F_X(y+d)$$

Espérance

$$E\left[Y^L\right] = E\left[(X-d)_+\right] = E[X] - E[X \wedge d] \tag{8.3}$$

$$E\left[Y^P\right] = \frac{E\left[(X-d)_+\right]}{S_X(d)} = \frac{E[X] - E[X \wedge d]}{S_X(d)} \tag{8.4}$$

1. Il est à noter que la fonction de Hazard n'est pas définie à 0.

Cette espérance s'appelle la prime *Stop-Loss*, et elle est définie par

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

8.2.2 Franchise deductible

Définitions

Lorsque la perte dépasse le déductible franchise de montant d, l'assureur assume l'entièreté des coûts [2]. Pour la v.a. Y^L, on a

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

Pour la v.a. Y^P,

$$Y^P = \begin{cases} \text{non-défini} & X \leq d \\ X & ,X > d \end{cases} \tag{8.5}$$

Fonctions reliées

Les fonctions de la v.a. Y^L sont

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

2. On voit plus souvent ce type de déductible dans un contexte d'invalidité : si on est absent plus d'un certain nombre de jours du travail, on se fait rembourser toutes ses absences en salaire.

$$F_{Y^L}(y) = \begin{cases} F_X(d) & ,0 < y \leq d \\ F_X(y) & ,y \geq d \end{cases}$$

$$S_{Y^L}(y) = \begin{cases} S_X(d) & ,0 < y \leq d \\ S_X(x) & ,y > d \end{cases}$$

$$h_{Y^L}(y) = \begin{cases} h_X(d) & ,0 < y \leq d \\ h_X(x) & ,y > d \end{cases}$$

Pour la fonction Y^P (*per-payment*),

$$f_{Y^P}(y) = \frac{f_X(d)}{S_X(d)}$$

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

$$S_{Y^P}(y) = \begin{cases} 1 & ,0 \leq y \leq d \\ \frac{S_X(y)}{S_X(d)} & ,y > d \end{cases}$$

$$h_{Y^P}(y) = \begin{cases} 0 & ,0 < y \leq d \\ h_X(y) & ,y > d \end{cases}$$

Espérance

$$E\left[Y^L\right] = E\left[(X-d)_+^F\right] = E[X] - E[X \wedge d] + dS_X(d) \tag{8.6}$$

$$E\left[Y^P\right] = E\left[(X-d)_+^F\right] = \frac{E[X] - E[X \wedge d]}{S_X(d)} + d \tag{8.7}$$

8.3 Loss Elimination Ratio

Définition 8.3.1 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 :

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

Mais on sait que

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

Alors,

$$LER = \frac{E[X \wedge d]}{E[X]} \tag{8.8}$$

Note sur l'inflation Il arrive en exercice qu'on nous demande de comparer ce ratio avec et sans inflation. Soit un contrat avec r % d'inflation. Alors, on peut trouver $E[(X-d)_+]$ qui tient compte de cette inflation :

$$E\left[(X-d)_+^I\right] = (1+r)\left(E[X] - E\left[X \wedge \frac{d}{1_r}\right]\right)$$

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,

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

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

□

8.4 Policy Limits

8.4.1 Définitions

Soit un contrat d'assurance où il est conclu que l'assureur débourse un maximum de u dans le montant de la perte. Ce type de modification au contrat créé une v.a. *censurée à droite*, i.e. le montant déboursé est maximisé à u. On définit Y comme étant

$$Y = \begin{cases} X & ,X \leq u \\ u & ,Y > u \end{cases} \tag{8.9}$$

8.4.2 Fonctions reliées

On peut déduire les fonctions reliées suivantes :

$$f_Y(y) = \begin{cases} f_X(y) & ,y < u \\ S_X(u) & ,y = u \end{cases}$$

$$F_Y(y) = \begin{cases} F_X(y) & ,y \leq u \\ 1 & ,y > u \end{cases}$$

Note sur l'inflation Si on a r % d'inflation, alors

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

Démonstration.

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

□

8.5 Coassurance, déductible et limites

Définitions

Le livre nous donne une fonction de perte qui englobe 4 éléments en même temps : l'inflation, les déductibles, la coassurance [3] et les limites de police. Pour la v.a. Y^L,

$$Y^L = \begin{cases} 0 & , x < \frac{d}{1+r} \\ \alpha\Big((1+r)x - d\Big) & , \frac{d}{1+r} \leq x < \frac{u}{1+r} \\ \alpha(u-d) & , x \geq \frac{u}{1+r} \end{cases} \tag{8.10}$$

et pour Y^P,

$$Y^L = \begin{cases} \text{Non-défini} & , x < \frac{d}{1+r} \\ \alpha\Big((1+r)x - d\Big) & , \frac{d}{1+r} \leq x < \frac{u}{1+r} \\ \alpha(u-d) & , x \geq \frac{u}{1+r} \end{cases} \tag{8.11}$$

Espérance

$$E\left[Y^L\right] = \alpha(1+r)\left(E\left[X \wedge \frac{u}{1+r}\right] - E\left[X \wedge \frac{d}{1+r}\right]\right) \tag{8.12}$$

$$E\left[Y^P\right] = \frac{E\left[Y^L\right]}{S_X\left(\frac{d}{1+r}\right)} \tag{8.13}$$

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

3. Dans ce type de contrat, la compagnie paie une proportion α de la perte, et le $(1-\alpha)$ restant est assumé par le titulaire de la police.

Chapitre 9

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

9.1 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

9.2 Estimation de données complètes

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

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} 1_{[x_i \leq t]} \tag{9.1}$$

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

$$F_n(t) = \begin{cases} 0 & , t < x_{(1)} \\ \frac{1}{n} & , x_{(1)} \leq t < x_{(2)} \\ \frac{2}{n} & , x_{(2)} \leq t < x_{(3)} \\ \ldots & \\ 1 & , t \geq x_{(n)} \end{cases} \tag{9.2}$$

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,

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

9.2.1 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 $_{i+1}$.

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

La formule est

$$F_n^{\text{OGIVE}}(x) = \frac{c_j - x}{c_j - c_{j-1}} F_n(c_j - 1) + \frac{x - c_{j-1}}{c_j - c_{j-1}} F_n(c_j) \tag{9.3}$$

Remarques

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

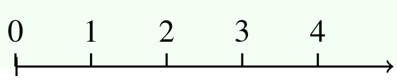
$$F_n(c_{j-1}) = F_n^{\text{OGIVE}}(c_{j-1})$$

(2) Si $x = c_j$,

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

Exemple 9.2.1 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 (9.3), on obtient

$$\begin{aligned}
f_n(x) &= \frac{\partial}{\partial x} F_x(x) \\
&= \frac{1}{c_j - c_{j-1}} F_x(c_j) - \frac{1}{c_j - c_{j-1}} F_n(c_{j-1}) \\
&= \frac{F_n(c_j) - F_n(c_{j-1})}{c_j - c_{j-1}}
\end{aligned} \tag{9.4}$$

9.3 Estimation de la fonction de survie

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

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

9.4 Estimateur Kaplan-Meier

9.5 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,

$$S(t) = \prod_{j \le t} p_j \tag{9.5}$$

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

$$\hat{S}(t) = \prod_{j \le t}(1 - \hat{q}_j) \tag{9.6}$$

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

9.5.1 Vu en classe 5 oct.

$$\begin{aligned}
\hat{S}(t) &= \prod_{i\leq t}\left(1-\frac{s_i}{r_i}\right)\\
S(t) &= e^{-H(t)}
\end{aligned}$$

ET

$$\begin{aligned}
\hat{H}(t) &= -\ln(\hat{S}(t))\\
&= \sum_{i\leq t} -\ln\left(1-\frac{s_i}{r_i}\right)\\
\hat{H}_{N-A}(t) & \sum_{i\leq t}\frac{s_i}{r_i}
\end{aligned}$$

où r_i est le nombre de survivant au temps i.

Exemple 9.5.1

Exemple fait en classe au début du cours, ligne du temps à insérer ici.