

Chapitre 3 : Méthodes Statistiques pour Analyse de Survie :

Méthodes Non-Paramétriques :

Ces méthodes d'estimation n'assument aucune distribution spécifique. Elles sont simples et fournissent des calculs utiles pour abréger les données de survie et faire des comparaisons par exemple, mais pour des conditions complexes, ces méthodes manquent de pertinence. Les modèles non paramétriques sont souvent préférés pour leur manque de restriction comparés à ceux paramétriques. Ils supposent moins d'hypothèses à-propos des données observées.

Estimateur de Kaplan Meier (KM) :

L'estimateur de Kaplan – Meier, également connu sous le nom d'estimateur de produit-limite (PL), est une méthode statistique non paramétrique utilisée pour estimer la fonction de survie à partir des données sur la durée de vie. Dans la recherche médicale, il est souvent utilisé pour mesurer la fraction de patients vivants pendant un certain temps après le traitement. Dans d'autres domaines, l'estimateur de Kaplan-Meier peut être utilisé pour mesurer la durée pendant laquelle les personnes restent au chômage après une perte d'emploi, le temps de défaillance des pièces de machines ou la durée pendant laquelle les fruits charnus restent sur les plantes avant d'être enlevés par les frugivores.

Kaplan-Meier calcule la probabilité de survie dans un petit intervalle de temps. Comme les données peuvent être censurées et tronquées en raison d'observations incomplètes, nous ne pouvons pas éliminer ces individus, tant qu'ils n'ont pas encore expérimenté l'événement, ils contribuent toujours à l'information du calcul et aussi, nous ne voulons pas réduire la taille de notre échantillon en les excluant. Car ce faire conduira également à un estimateur biaisé. Kaplan-Meier est considéré comme le moyen le plus simple d'estimer les probabilités de survie pour les durées de survie censurées et non censurées.

Nous calculons la probabilité de survie à des moments distincts en divisant le nombre de sujets survivants par le nombre de sujets au risque. Les sujets censurés ne sont pas considérés comme au risque, ils ne sont donc pas ajoutés au dénominateur n . Mathématiquement, nous pouvons exprimer la probabilité de survie estimée à un certain moment par la formule :

$$1 - \frac{d}{n}$$

Où,

d = nombre des sujets morts / nombre d'occurrence de l'événements

n = nombre des sujet vivant au début de la période concerné / les sujets au risque

La probabilité totale de survie ou probabilité cumulée dans la période de suivi est obtenue en multipliant toutes les probabilités de survie à tous les moments spécifiques dans un intervalle spécifié.

Pour mieux expliquer, notons les temps distincts d'occurrence de l'événement $t_1 < t_2 < t_3 \leq \dots \leq t_j$, où j patients ont expérimenté l'événement durant l'étude. Au temps t_1 la probabilité est p_1 , a temps t_2 la probabilité est p_2 sachant que le patient a survécu au temps t_1 , et au t_j la probabilité est p_j ayant survécu le temps t_{j-1} .

La probabilité de survivre au-delà du temps t_j est estimée à :

$$\hat{S}(t_j) = \hat{P}(T > t_j) = p_1 \cdot p_2 \dots p_{t_j}$$

On trouve alors l'estimateur de Kaplan-Meier par la formule :

$$\hat{S}(t_k) = \prod_{t_k < t} S(t_{k-1}) \left(1 - \frac{d_k}{n_k}\right) \quad 1 < k < j \quad (2.6)$$

Dans l'équation (2.6) $S(t_{k-1})$ est la probabilité de survie calculée au temps t_{k-1} , d_k est le nombre des sujets décédés au temps t_k et n_k est le nombre des sujets vivants juste avant t_k . On convient que $S(0) = 1$.

Par exemple, la probabilité de survie d'un patient deux jours après une greffe de rein pourrait être trouvée comme le nombre de patients ayant survécu le premier jour $S(t_{k-1})$ multipliée par la probabilité qu'il survit le deuxième jour sachant qu'il a survécu le premier jour $(1 - \frac{d_k}{n_k})$. Cette deuxième probabilité est une probabilité conditionnelle, c'est-à-dire que pour qu'un individu reste au risque le deuxième jour il doit avoir survécu le premier.

Exemple. La figure 2.6 illustre la courbe de la fonction de survie en prenant des données hypothétiques d'un groupe de patients entrés dans un essai clinique recevant un traitement antirétroviral pour l'infection par le VIH.

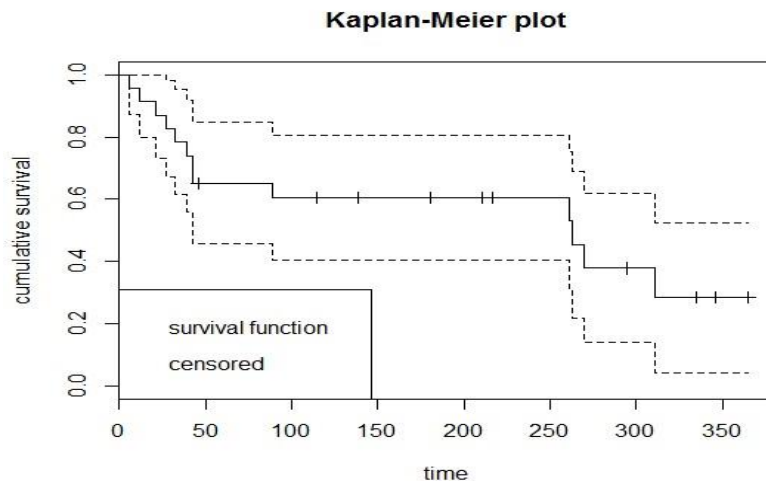


Figure 9: Courbe de l'estimateur Kaplan-Meier pour un groupe de patients recevant ARV thérapie

Les données temps à événement i.e., la mort, parmi les patients est :

6,12,21,27,32,39,43,43,46*,89,115*,139*,181*,211*,217*,261,263,270,295*,311,335*,346*,365* (* représente des données censures à droite).

À partir de la figure 9, nous pouvons voir que la probabilité estimée est une fonction en escalier qui reste inchangée durant une même période d'observation même s'il y en a une censurée. L'axe X (ligne horizontale) montre le temps passé après l'entrée dans les études et l'axe Y (ligne verticale) montre les probabilités de survie estimées. Le temps t où la probabilité cumulée est de 0,5, c'est-à-dire $S(t) = 0,5$ est appelé temps médian de survie qui, selon cet exemple, est $t=263$. Nous pouvons utiliser différents programmes statistiques pour tracer la courbe de Kaplan-Meier tels que SPSS, R, Sigma plot, python etc.

Comparaison des estimations Kaplan-Meier

Les courbes de Kaplan-Meier sont un bon moyen pour faire des comparaisons entre s données de survie. Par exemple, nous pouvons vérifier si un traitement particulier A, administré à un groupe de patients, est moins ou plus efficace que le nouveau traitement B administré à un autre groupe. Les modèles de survie dans les courbes de survie tels que les écarts horizontaux et verticaux peuvent être comparés. L'écart dans le sens horizontal signifie qu'un des deux groupes a mis plus de temps à vivre l'événement (la mort) et l'écart dans le sens vertical signifie qu'un groupe a survécu plus que l'autre groupe. Cela signifie que les deux directions sont deux côtés du même résultat, nous n'avons donc pas besoin de vérifier les deux directions en même temps. Dans les essais cliniques, la comparaison des courbes de survie est particulièrement prise en compte. La différence doit être statistiquement significative sinon les deux estimations sont considérées comme identiques.

Il existe plusieurs méthodes pour comparer les courbes de survies, le plus commun est le test log-rank. Ce test calcule le chi deux (χ^2) pour chaque temps d'occurrence de l'événement et fait la somme pour chaque groupe. Le chi deux final est obtenu en additionnant les résultats de chaque groupe pour comparer leurs groupes.

Log Rank Test

Dans cette méthode, nous comparons les courbes de deux groupes de patients différents et testons si leurs durées de survie sont statistiquement différentes ou non en utilisant un test statistique qui testant une hypothèse dite nulle. L'hypothèse nulle suppose qu'il n'y a pas de différence entre les courbes concernant la survie. Nous calculons les statistiques de test du log-rank comme suivant :

$$\chi^2 = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} \quad (2.7)$$

Où,

O_1 = nombre totale d'occurrence de l'événement (patients morts) dans le groupe 1

O_2 = nombre totale d'événement observés dans le groupe 2

E_1 = nombre d'événements attendus pour le groupe 1

E_2 = nombre d'événements attendus pour le groupe 2

Le nombre total d'événements attendus dans l'un des groupes est la somme de tous les événements attendus calculés à différents moments (au moment de chaque événement) et le nombre d'événements attendus au moment de chaque événement dans un groupe est calculé en multipliant le risque d'événement à ce moment avec le nombre de tous les patients vivants au début d'un événement dans ce groupe (c'est-à-dire si le nombre total des patients est 46, 23 dans chaque groupe, et au jour 6, le risque d'événement est calculé comme $1/46 = 0,0217395$ où tous les patients sont vivants au début de la journée et 1 est décédé, donc dans le groupe 2, le nombre attendu d'événements au jour 6 serait de $23 \times 0,0217395 = 0,5$). Une fois que nous obtenons la somme de tous les événements attendus dans le groupe 2 (E_2), nous pouvons obtenir (E_1) en soustrayant E_2 de $O_1 + O_2$.

Exemple :

Suivant l'exemple précédant de courbe de Kaplan-Meier, prenons d'autres données hypothétiques pour des patients entrés dans un test clinique pour recevoir une nouvelle thérapie ayurvédique contre l'infection par le VIH :

9,13,27,38,45*,49,49,79*,93,118*,118*,126,159*,211*,218,229*,263*,298*,301,333,346*,353*,362* (* signifie observation censurée à droite) et nommons-le groupe 2 (thérapie ayurvédique).

Figure 2.7 illustre la différence entre les courbes de survie des deux groupes.

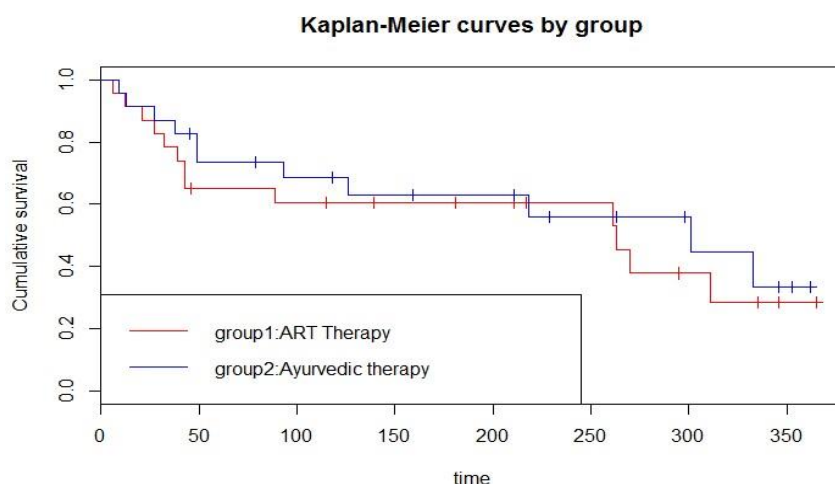


Figure 10: courbes d'estimations Kaplan-Meier de deux groupes de patients

Depuis la Figure 10, On constate qu'il n'y a pas une grande différence entre les deux courbes. Mais pour vérifier la différence avec précision, on effectue des tests statistiques (calculés par la formule de l'équation (2.7)) et on compare par le biais de la valeur critique (tirée depuis la table chi-deux) pour un degré de liberté. Si la Valeur du test statistique est inférieure à la valeur critique, on accepte l'hypothèse nulle. Pour l'exemple précédent, Table 1 montre le résultat de la formule appliquée pour le log rank test.

	N	Observées	Attendues	$\frac{(O-E)^2}{E}$	$\frac{(O-E)^2}{V}$
Groupe 1	23	13	11.5	0.184	0.357
Groupe 2	23	11	12.5	0.170	0.357
$\chi^2 = 0.4$ pour un degré de liberté, $p = 0.5$					

Table 1 : Résultat du log rank test par programmation statistique R

Le résultat, $\chi^2 = 0.4$ avec $p \text{ value} = 0.5$, montre que la $p \text{ value}$ est la même que le niveau significatif, ce qui veut dire qu'il n'y a pas de différence significative entre les deux groupes 1 et 2 (l'hypothèse nulle est acceptée). Les résultats généraux des deux thérapies sont identiques vis-à-vis la survie.

Estimateur de Nelson–Aalen :

L'estimateur de Nelson-Aalen est un estimateur non paramétrique de la fonction de taux de risque cumulé. Construit par Wayne Nelson et Odd Aalen. Il est utilisé dans l'analyse de survie afin d'estimer le nombre cumulé d'événements attendus. L'estimateur est donné par :

$$\tilde{H}(t) = \sum_{t_i < t} \frac{d_i}{n_i}$$

Où d_i est le nombre des événements produits au temps t_i , et n_i le nombre total des individus au risque au temps t_i .

Une courbe concave de l'estimateur Nelson-Aalen est un indice de la mortalité infantile tandis qu'une courbe convexe indique une mortalité par usure.

Cet estimateur peut être utilisé par exemple pour tester l'homogénéité des processus de Poisson

Tables de survie :

Cette méthode fait une subdivision de la période d'observation en des intervalles de temps plus petits. On utilise pour chaque intervalle, toutes les observations faites au moins sur cette durée dans le calcul de la probabilité de l'événement final qui se produit au cours de cet intervalle. Les probabilités estimées sur chaque intervalle sont ensuite utilisées pour estimer la probabilité globale de l'événement se produisant à différentes dates.

Appelées aussi tables de mortalité, ou encore analyse actuarielle car elle est basée sur des méthodologies fréquemment utilisées dans les cercles actuariels pour calculer les taux d'assurance en fonction des risques de mortalité et des paiements de prestations ultérieurs, la plupart de ces tables sont transversales ou actuelles. C'est-à-dire les individus suivis pour l'intervalle $[n; n + 1[$ ne sont pas les mêmes que ceux suivis pour l'intervalle $[n + 1; n + 2[$ et les deux groupes des deux intervalles de temps sont suivis simultanément. Au contraire des tables de survie longitudinale, où un même groupe est suivi sur une longue durée d'un intervalle de temps au suivant.

Les tables de survie fournissent une analyse facile à comprendre et facile à réaliser des données obtenues à des intervalles prédéterminés. Elles supposent que l'événement se produit de manière aléatoire tout au long de l'intervalle et, si la taille de l'échantillon est grande, se produira, en moyenne, au centre de chaque intervalle. On suppose aussi que la probabilité de survie à un intervalle, bien que conditionnelle à la survie des intervalles précédents, est indépendante de la probabilité de survie à l'intervalle ou aux intervalles précédents.

Méthodes semi-Paramétriques :

Cox Régression :

Le modèle de Cox est un modèle semi paramétrique largement utilisé dans l'analyse de survie, que les données censurées soient de temps à événement continu ou discret.

Hypothèse des risques proportionnels :

La propriété principale du modèle Cox est que les deux fonctions de risque $h_1(t)$ et $h_0(t)$ pour deux distributions indépendantes soient proportionnelles :

$$h_1(t) = \psi^x h_0(t), \forall t > 0, x = 0; 1, \psi > 0 \quad (2.8)$$

Où ψ est la constante positive de proportionnalité indépendante de t et $h_0(t)$ est le risque de base.

L'hypothèse des risques proportionnels n'est pas toujours vérifiée et doit toujours être soigneusement examinée. Cela pourrait être fait en utilisant les résidus de Schoenfeld.

Model de Cox aux risques proportionnels :

Il est intéressant de savoir si les attributs d'un individu sont associés à l'événement d'intérêt. Si c'est le cas, ces caractéristiques/attributs sont appelés covariables ou facteurs de risque. L'effet de ces facteurs sur le temps jusqu'à l'événement peut être modélisé par le modèle de Cox. D'autre part, le hasard est la probabilité de vivre l'événement étant donné que les patients ont survécu à une certaine période de temps.

Le modèle de Cox est un modèle de régression pour les données d'analyse de survie qui supposant que les covariables affecteront les temps de survie. Il permet de tester la différence entre les durées de survie de différents groupes d'individus, ce qui permet de prendre en compte d'autres facteurs (c'est-à-dire des

covariables). L'hypothèse de risque proportionnel est la base de ce modèle. En utilisant $\beta = \log(\psi)$ on peut réécrire l'équation (2.8), pour estimer la fonction de risque comme :

$$h_x(t) = h_0(t)e^{\beta x}, \quad t > 0, x = 0; 1, -\infty < \beta < +\infty \quad (2.9)$$

Qui est la forme du modèle de Cox de deux groupes, où t est le temps de survie, $h_x(t)$ est le hasard au temps t , β est le paramètre à estimer, $h_0(t)$ est le hasard de base (risque quand tous les covariables sont égales à zéro) et x est la covariable (également appelée variables explicatives). C'est à cause de ces deux termes β et $h_0(t)$, que le modèle de Cox est appelé un modèle semi-paramétrique où $h_0(t)$ est non paramétrique alors que β est une partie paramétrique.

Pour comprendre le paramètre β on doit remarquer que la fonction de risque est multipliée par e^β à chaque fois que la variable x augmente d'une unité. Par exemple si pour deux groupes A et B, la variable x prend les valeurs 0 et 1, on dit que le groupe B a un risque e^β fois plus grand que le groupe A.

$$e^{\beta x} = \frac{h_1(t)}{h_0(t)}, \quad \forall t \geq 0 \quad (2.10)$$

L'équation (2.10) est appelée taux de risque (le risque d'échec).

Le modèle général du risque proportionnel pour un ensemble de p covariables $x = (x_{1i}; x_{2i}; \dots; x_{pi})^T$, est donné sous la forme :

$$h_i(t, x_i) = h_0(t)e^{\beta^T x_i}, \quad t > 0 \quad (2.11)$$

Où $\beta = (\beta_{1i}; \beta_{2i}; \dots; \beta_{pi})^T$ $i = 1, 2, \dots, d$, est le coefficient de régression. $h_0(t)$, le risque de base est le risque quand toutes les covariables sont égales à zéro $(x_{1i}; x_{2i}; \dots; x_{pi}) = (0; 0; \dots; 0)$. Si nous avons deux patients avec le même score sur toutes les covariables sauf la covariable m alors :

$$e^{\beta_m} = \frac{h_1(t)}{h_0(t)}, \quad \forall t \geq 0$$

L'effet de la covariable x_m peut être interprété comme si x_m augmente d'une unité, alors le risque est multiplié par e^{β_m} .

Estimation

Dans l'équation (2.11), deux composantes doivent être estimées. Le plus important en est le coefficient de régression β puis le risque de base $h_0(t)$.

Dans le modèle Cox à risque proportionnel, le paramètre β peut être estimé par la vraisemblance partielle. Or celle-ci ne peut pas être utilisée si on ne connaît rien sur le risque de base $h_0(t)$ car elle n'aurait aucune forme spécifique. Aussi ne prend-elle pas en considération la distribution des valeurs censurées qui sont éliminées de la formule de Cox, d'où le nom « vraisemblance partielle ». Elle ne prend donc en considération que les sujets ayant vécu l'événement. Le paramètre de régression pour le modèle Cox est obtenu en maximisant la vraisemblance partielle.

Supposant que $t_{(i)} = t_{(1)}, t_{(2)}, \dots, t_{(d)}$ soient les vrais temps d'occurrence de l'événement avec une occurrence à chaque temps, et $R(t_{(i)})$ est le risque du groupe d'individus observés i.e. qui ne sont ni censurés ni échoués (n'ont pas encore eu l'événement) au temps t_i , $i = 1, 2, \dots, d$. Alors la vraisemblance totale est :

$$L(\beta) = \prod_{i=1}^k L_i(\beta) = \prod_{i=1}^k P(\text{No. } i \text{ meurt} | \text{une occurrence d'événement } R_i)$$

$$L_i(\beta) = \frac{h_i(t_i)}{\sum_{l \in R_i} h_l(t_i)}$$

En utilisant l'équation (2.9)

$$L_i(\beta) = \frac{h_0(t_{(i)})e^{\beta^T x_i}}{\sum_{l \in R_i} h_0(t_{(i)})e^{\beta^T x_l}}$$

On peut simplifier le risque de base du dénominateur et du numérateur pour obtenir la forme finale de la vraisemblance partielle :

$$L(\beta) = \prod_{i=1}^k L_i(\beta) = \prod_{i=1}^k \frac{e^{\beta^T x_i}}{\sum_{l \in R_i} e^{\beta^T x_l}} \quad (2.12)$$

L'estimation du maximum de vraisemblance partielle de β peut être obtenue en calculant le log de la vraisemblance partielle :

$$l(\beta) = \log(L(\beta)) = \log\left(\prod_{i=1}^k \frac{e^{\beta^T x_i}}{\sum_{l \in R_i} e^{\beta^T x_l}}\right)$$

$$l(\beta) = \sum_{i=1}^k [\beta^T x_i - \log(\sum_{l \in R_i} e^{\beta^T x_l})] \quad (2.13)$$

La fonction de Score $U(\beta)$ est définie comme la dérivée première du log de vraisemblance :

$$U(\beta) = \frac{\partial}{\partial \beta} (l(\beta)) = x_i - \frac{\sum_{l \in R_i} x_l e^{\beta^T x_l}}{\sum_{l \in R_i} e^{\beta^T x_l}} \quad (2.14)$$

On tire l'estimation $\hat{\beta}$ du paramètre β , par :

$$U(\hat{\beta}) = 0$$

$$x_i - \frac{\sum_{l \in R_i} x_l e^{\beta^T x_l}}{\sum_{l \in R_i} e^{\beta^T x_l}} = 0 \quad (2.15)$$

En prenant le négatif de la dérivée de la fonction de score (ou dérivée seconde du log de vraisemblance) on peut trouver la matrice d'information de la vraisemblance partielle observée $I(\beta)$.

$$I(\beta) = -\frac{\partial}{\partial \beta} U(\beta) = -\frac{\partial}{\partial \beta} \left(x_i - \frac{\sum_{l \in R_i} x_l e^{\beta^T x_l}}{\sum_{l \in R_i} e^{\beta^T x_l}} \right) \quad (2.16)$$

Après avoir obtenu la vraisemblance partielle maximale, $\hat{\beta}$. Alors asymptotiquement,

$$\hat{\beta} \sim N(\beta_0, I^{-1}(\hat{\beta}))$$

Où, $I^{-1}(\hat{\beta})$ est la matrice inverse de la matrice d'information à $\beta = \hat{\beta}$ et β_0 est la valeur exacte. Cette distribution approximative est utilisée pour construire un intervalle de confiance et tester l'hypothèse $H_0: \beta = \beta_0$.

Le risque de base peut être estimé en utilisant la fonction de risque cumulatif.

Soit d_j le nombre d'événements et R_j le risque à t_j . L'estimateur est :

$$\hat{H}_0(t) = \sum_{j, t_j \leq t} \frac{d_j}{\sum_{l \in R_j} e^{\hat{\beta} x_l}} \quad (2.17)$$

Et si $\hat{\beta} = 0$, l'équation (2.17) devient :

$$\hat{H}_0(t) = \sum_{j, t_j \leq t} \frac{d_j}{n_j}$$

Où n_j est la taille à R_j .

Résidus de Schoenfeld

Pour le modèle de régression à risque proportionnel, Schoenfeld a recommandé une statistique de qualité de l'ajustement du chi deux qui exploite les résidus de la forme « attendu-observé ». Les résidus de Schoenfeld sont définis comme : $r_k(\hat{\beta}) = X_{(k)} - \bar{x}(\beta, t_k) \quad k = 1, \dots, d$

Où d est le nombre total des événements, $X_{(k)}$ est l'individu qui a eu l'événement k au temps t_k . et $\bar{x}(\beta, t_k)$ est la moyenne pondérée de X .

Méthodes paramétriques :

On suppose que la distribution des durées de survie appartient une famille de loi paramétrique donnée. Ainsi, le modèle paramétrique peut être formulé en précisant la forme de l'une ou l'autre des cinq fonctions Équivalentes qui définissent la loi de la durée : λ , Λ , f , S ou F . Néanmoins, on spécifie souvent la forme du risque instantané λ : constant, monotone croissant ou décroissant et en forme de \cap ou de \cup .

Les estimateurs des paramètres du modèle sont ensuite obtenus en maximisant la vraisemblance des observations (par l'intermédiaire de méthodes itératives, par exemple l'algorithme de Newton-Raphson).

Risque instantané constant (loi exponentielle) :

La loi exponentielle $\varepsilon(\theta)$, qui ne dépend que d'un paramètre θ , est la seule qui admet un risque instantané constant. Cette loi est aussi dite "sans mémoire" car la probabilité de décès pour un individu dans un certain laps de temps est la même quelle que soit sa durée de vie (i.e. $P(X > t + s | X > t) = P(X > s)$). Les quantités associées cette loi sont :

$$\begin{aligned} f(t/\theta) &= \theta e^{-\theta t}, \\ \lambda(t/\theta) &= \theta, \\ S(t/\theta) &= e^{-\theta t} \end{aligned}$$

Dans certaines applications, on peut découper le temps en plusieurs intervalles et considérer un θ_i différent pour chacun des intervalles (risque est constant sur chaque période mais varie d'une période une autre).

Risque instantané monotone :

Loi de Weibull :

Considérons une loi de Weibull $W(\theta, \nu)$, alors

$$f(t | \theta, \nu) = \nu \left(\frac{1}{\theta}\right)^\nu t^{\nu-1} \exp\left(-\left(\frac{t}{\theta}\right)^\nu\right), \quad t \geq 0 \text{ et } \theta, \nu > 0$$

$$\lambda(t | \theta, \nu) = \nu \left(\frac{1}{\theta}\right)^\nu t^{\nu-1},$$

$$S(t | \theta, \nu) = \exp\left(-\left(\frac{t}{\theta}\right)^\nu\right).$$

Pour $\nu = 1$, on retrouve la loi exponentielle $\varepsilon(\frac{1}{\theta})$. Si $0 < \nu < 1$, le risque instantané est monotone décroissant ; si $\nu > 1$ le risque est monotone croissant.

Loi Gamma

Considérons une loi Gamma $G(\nu, \theta)$, alors

$$f(t | \nu, \theta) = \frac{\theta^\nu}{\Gamma(\nu)} t^{\nu-1} e^{-\theta t}, \quad t \geq 0 \text{ et } \theta, \nu > 0$$

$$F(t | \nu, \theta) = \frac{1}{\Gamma(\nu)} \int_0^{\theta t} u^{\nu-1} e^{-u} du,$$

$$\lambda(t | \nu, \theta) = \frac{f(t | \theta, \nu)}{1 - F(t | \theta, \nu)},$$

Avec $\Gamma(\nu) = \int_0^\infty u^{\nu-1} e^{-u} du$ Pour $\nu = 1$, on retrouve la loi exponentielle $\varepsilon(\theta)$. Si $0 < \nu < 1$, le risque instantané est décroissant ; si $\nu > 1$ le risque est croissant

Risque instantané en \cap et \cup :

Lois de Weibull généralisées :

La loi de Weibull est intéressante pour modéliser des risques monotones. Cependant elle devient mal adaptée quand les risques sont en forme de cloche. Une alternative est l'utilisation de la loi de Weibull généralisée $GW(\theta, \nu, \gamma)$:

$$\lambda(t/\theta, \nu, \gamma) = (1 + (t/\theta)^\nu)^{\frac{1}{\gamma}-1} \frac{\nu}{\gamma \theta^\nu} t^{\nu-1}, \quad t \geq 0, \theta, \nu, \gamma > 0,$$

$$S(t/\theta, \nu, \gamma) = \exp [1 - (1 + (t/\theta)^\nu)^{\frac{1}{\gamma}}]$$

Pour $\gamma = 1$, on retrouve la loi de Weibull $W(\theta, \nu)$; pour $\gamma = 1$ et $\nu = 1$, on retrouve la loi exponentielle $\varepsilon(\frac{1}{\theta})$. En faisant varier les paramètres on peut obtenir des risques constants, monotones croissants et décroissants, avec des formes en \cap (par exemple, $\theta = 10, \nu = 0.5, \gamma = 4$) et des formes en \cup (par exemple, $\theta = 130, \nu = 5, \gamma = 0.1$).

Les risques en forme de cloche sont souvent présents dans le domaine du vivant. Par exemple, les risques instantanés en forme de \cup comportent 3 phases : la période de mortalité infantile, la période risque faible et la période de vieillissement durant laquelle le risque augmente.

La loi Log-logistique de paramètre (α, β) :

$$F(t) = \frac{t^\beta}{\alpha^\beta + t^\beta}, \quad \alpha > 0, \beta > 1$$

La Log-normales de paramètre (μ, σ) :

$$F(t) = \Phi\left(\frac{\ln t - \mu}{\sigma}\right), \quad \sigma > 0, t > 0$$

La loi gaussienne inverse (μ, λ) :

$$f(t) = \sqrt{\frac{\lambda}{2\pi t^3}} \exp\left(-\frac{\lambda}{2\mu^2 t}(t - \mu)^2\right), \quad \lambda > 0, \quad t \geq 0$$

Gompert-Makeham :

Loi modélisant la mortalité humaine, modèle de référence pour la construction de tables de mortalité.

$$h(t) = a + bc^t$$

$$S(t) = \exp\left(-at - \frac{b}{\ln(c)}(c^t - 1)\right)$$

Où,

a , est le taux de décès accidentel (indépendant de l'âge).

bc^t est le vieillissement exponentiel (si $c > 1$).