ANALYSE DE LA SURVIE

Une étude de survie est une étude :
→ longitudinale (suivi des personnes au cours du temps)
→ prospective (prise en compte des événements survenant dans la durée de l'étude)
→ de cohorte (observation d'un groupe de personnes dans le temps)

♡ L'analyse de la survie est **l'estimation de la probabilité de survenue d'un événement** (décès*, complication post opératoire, rechute, guérison) dans le temps, en fonction de facteurs pronostiques
→ Probabilité de survivre au moins un certain temps « t » à compter d'un instant de référence
→ Probabilité pour que l'évènement attendu survienne après un certain délai
♡ L'analyse de la survie est aussi **l'étude comparative de la survenue dans le temps d'un événement dans différents groupes**

L'événement considéré doit être défini de la même manière pour tous les sujets

I. Définitions

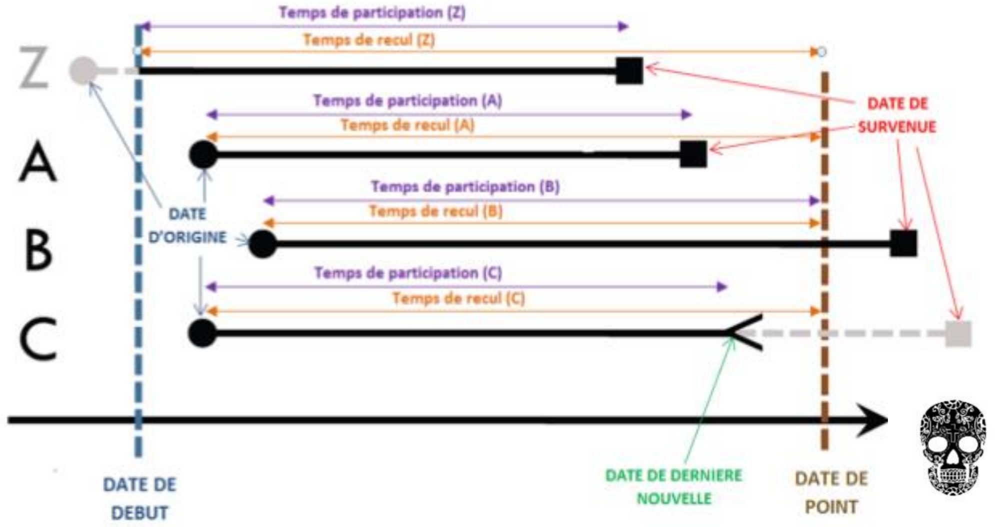
- Facteur pronostic : facteur susceptible d'expliquer la survenue du décès (ou d'un autre événement) au cours du temps

/!\ *Un facteur pronostique ~~est un évènement~~*
Un facteur pronostique ~~est un facteur de risque~~
Un facteur pronostique ~~explique la survenue d'une maladie~~

- Cohorte : ensemble de sujets qui vivent les mêmes événements au même moment

☮ Cohorte « Incipiente » : Dans ce cas, la cohorte des patients qui rentrent dans l'étude doit inclure des sujets observés au début de leur affection au même stade de leur maladie (« cas incidents »)

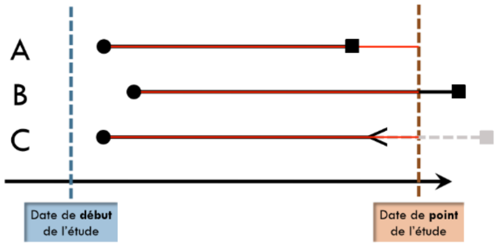
- Evénement d'intérêt : événement auquel on s'intéresse au cours de l'étude *(l'événement d'intérêt n'est pas forcément le décès, mais peut être aussi la survenue d'une maladie, la récurrence de symptômes après traitement)*

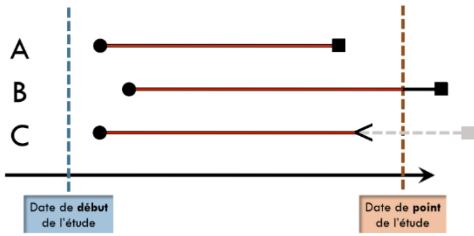


Les dates	Date d'origine	Date de point	Date des dernières nouvelles
Définitions	C'est la date correspondant au point de départ de la surveillance	C'est la date choisie pour faire le bilan	C'est la date la plus récente à laquelle on a recueilli des informations sur le patient *(notamment la survenue ou non de l'événement d'intérêt)*
Particularités	Peut être **différente** pour chaque sujet selon les modalités d'inclusion du sujet *Date d'origine peut être antérieure à l'inclusion dans l'étude : cohorte historique*	Au-delà cette date les informations recueillies **ne sont plus** considérées dans l'analyse	

- Perdu de vue : Un sujet est perdu de vue lorsque sa surveillance est interrompue avant la date de point et que l'événement d'intérêt ne s'est pas produit
- Censure : Une durée de survie d'un individu est dite censurée lorsque l'événement d'intérêt n'a pas été observé. Elle concerne : les sujets perdus de vus (C) et ceux vivant à la date de point (B)

Les temps	**Temps de recul**	**Temps de participation**
Définitions	Délai entre la date d'origine et la date de point, c'est-à-dire le délai maximum potentiel de suivi pour un sujet	Durée de surveillance pour chaque sujet utilisée dans l'estimation de la survie
Particularités	Les reculs minimum et maximum d'une série de sujets définissent donc l'ancienneté de cette série	3 situations : → L'événement a lieu **au cours** de la surveillance → Le sujet est **vivant** à la date de point → Le sujet est **perdu de vue**





II. Fonction de survie

A) Loi exponentielle

Rappel : **La loi de Poisson** régit la survenue d'un événement par unité de mesure (temps, volume, surface…)

Si un événement se réalise selon une loi de Poisson (de paramètre λ=μ=σ2), le temps entre deux réalisations consécutives de l'événement considéré est distribué selon une loi exponentielle d'espérance 1/λ

λ : taux de défaillance instantanée

→ Utilisée couramment pour représenter la **durée de vie** de composants ou d'équipements pour lesquels on suppose que le taux de défaillance "λ" est constant au cours du temps (la durée de vie au-delà de « t » est indépendante de « t »)

Fonction de densité de la loi exponentielle : $\boxed{f(x) = \lambda e^{-\lambda x}}$

Fonction de répartition de la loi exponentielle : $\boxed{F(t) = P(X <= t) = 1 - e^{-\lambda t}}$

F(t) représente la proportion d'équipement qui tombent en panne avant le temps « t »

La quantité **1-F(t)** représente donc la quantité d'équipements qui fonctionnent pendant une durée au moins égale à t cette quantité se note **S(t)** et s'appelle la **fonction de survie** ♡

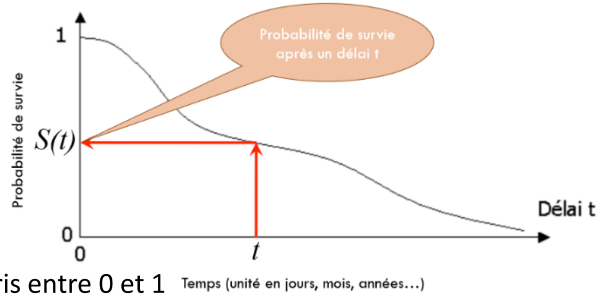
B) Fonction de survie

C'est la probabilité pour que l'événement d'intérêt « T » (le décès par exemple) intervienne après un délai supérieur à « t ». Autrement dit, que l'événement d'intérêt « T » ne survienne pas avant la date « t »

$$\boxed{S(t) = 1\text{-}F(t) = P(X>t) = e^{-\lambda t}}$$

La fonction de survie est représentée graphiquement par une **courbe de survie**

C) La courbe de survie



C'est une fonction :

→ Décroissante

→ Avec en ordonnée S(t) compris entre 0 et 1

D) Calcul de probabilité

On rappelle que l'on note T l'évènement d'intérêt

- La fonction de survie permet de calculer la probabilité pour que le décès survienne après un délai t1 et avant un délai t2

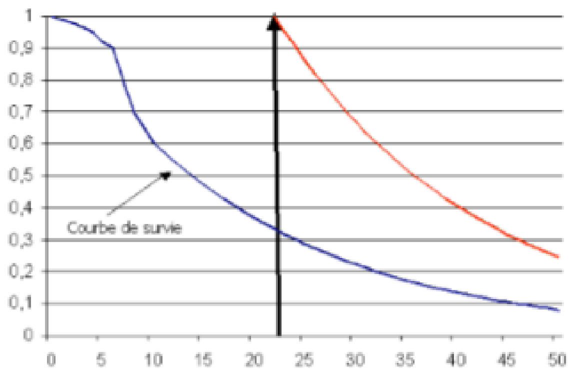
$$\boxed{\mathbf{P}(\mathbf{T}\epsilon[\,\mathbf{t1}\,;\mathbf{t2}\,]) = \mathbf{S(t1)} - \mathbf{S(t2)}}$$

- Elle donne aussi la probabilité de survivre encore après un délai « t » sachant que l'on est survivant après un délai « τ » = probabilité conditionnelle = S(t/τ). On a : S(t) = P(T>t) et S(τ) = P(T> τ)

$$\boxed{\mathbf{S(t\ sachant\ \tau)} = \mathbf{S(t|\tau)} = \frac{\mathbf{S(t)}}{\mathbf{S(\tau)}}}$$

Je vous ai perdu ? Un petit exemple et c'est reparti : ☺

Exercice : Supposons que l'on veuille calculer la probabilité de survivre après un délai t =33ans sachant que l'on ait vivant à τ = 23 ans



La courbe bleue décrit la probabilité de survie après t

→ Proportion de **survivant à 23 ans** = 33% = probabilité de survivre après 23 ans = S(t) = P(T>23)

→ Proportion de **survivant à 33 ans** = 20% =probabilité de survivre après 33 ans = S(t) = P(T>33)

La courbe rouge : décrit la probabilité de survivre après « t » **sachant** que l'on est vivant à 23 ans

→ Proportion de survivant à 23 ans = 100% : Il s'agit de la probabilité de **survivre après 23 ans sachant que l'on est vivant à 23 ans** : S(23/23) = $\frac{0,33}{0,33}$ = 1

→ Proportion de survivant à 33 ans = 60% : Il s'agit de la probabilité de **survivre après 33 ans sachant que l'on est vivant à 23 ans :** S(33/23) = $\frac{S(33)}{S(23)} = \frac{0,2}{0,33}$ = 0,6

III. Estimation de la survie I Will Survive

A) Recueil des données

On s'intéresse à : *(petits rappels)*

- ✓ **La date d'origine** : la date à laquelle a débuté l'observation

Exemple : la date de diagnostic du cancer broncho-pulmonaire

- ✓ **La date des dernières nouvelles** : **la date de décès** pour les patients décédés **ou** la date à laquelle on dispose **des dernières données** relatives à l'état du patient sachant qu'il n'est pas décédé
- ✓ **La date de point** : la date à laquelle on fait le point ou date de fin d'observation :

/! Tout patient chez qui l'événement d'intérêt n'a pas été observé à la date de point est **censuré** à cette date. Un **sujet perdu de vue** à la date de point sera censuré à la date de dernières nouvelles

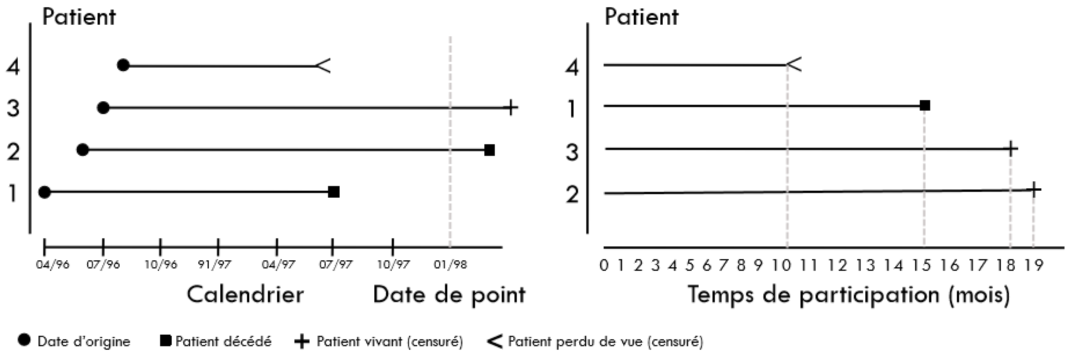
- ✓ **Un événement binaire « Tout ou rien »** : correspondant à l'état du patient en deux éventualités (vivant ou décédé) à la date des dernières nouvelles

B) Calcul des durées de suivi

♡À partir de ces données, les durées de suivi (ou temps de participation) de chaque patient sont calculées par **différence**

→ Délai entre la date d'origine et la date des dernières nouvelles

Mais la **date des dernières nouvelles** peut être :
→ La date de décès en cas de décès
→ La date de point pour les patients vivants pour lesquels le suivi est assuré
→ La date de perte de vue pour les patients vivants n'étant plus suivi dans la cohorte à la date de point



C) Calcul de la survie

*Dans le plus idéal des mondes où **aucune** variables n'est **censurée** (ne se produit jamais car un certain nombre de sujets seront perdus de vue, et un certain nombre seront encore vivants à la date de point) : **la fonction de survie se calcule par le pourcentage de survivants en fonction du temps***

Les 2 méthodes d'analyse non paramétriques de préférence utilisées sont :
→ **La méthode Actuarielle** : Utilisée lorsque les échantillons sont grands **> 200** sujet
→ **La méthode de Kaplan-Meier** : Utilisée lorsque les échantillons sont **< 200** sujets

Ces deux méthodes supposent que les probabilités de survie sont indépendantes du calendrier. *Exemple : la survie à 1 an d'un groupe de patients inclus en 1970 est identique à celle d'un groupe de patients inclus en 1990*

Méthode Actuarielle (n > 200)	**Méthode de Kaplan-Meier (n < 200)**
La fonction de survie est calculée sur des intervalles de temps fixés à priori (mois, trimestre, année)	Les intervalles sont définis par les instants auxquels les évènements sont observés → Ces intervalles sont donc inégaux, débutent à l'instant d'un événement et s'arrêtent juste avant l'événement suivant.

Pour chaque intervalle de temps on définit :
- **V** : Nombre de **sujets vivants** au début de l'intervalle :
- **D** : Nombre de **sujets décédés** dans l'intervalle :
- **C** : Nombre de **sujets vivants aux dernières nouvelles**, dont le temps de participation s'arrête dans l'intervalle = **censure**

N : Nombre de sujet exposés au risque d'événement sur l'intervalle

Méthode Actuarielle	Méthode de Kaplan-Meier
$N = V - \frac{C}{2}$	$N = V - C$

Probabilité d'événement durant l'intervalle : $\frac{D}{N}$

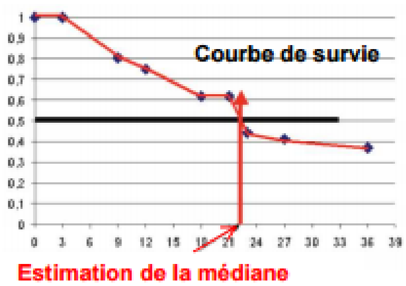
Probabilité de survie = survie instantanée : $\frac{N-D}{N}$

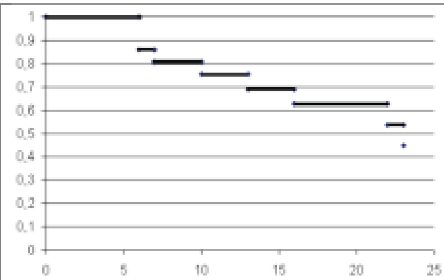
La fonction de survie $S(t)$ est estimée en faisant le **produit des survies instantanées** calculées sur tous les intervalles. *Exemple : Survie à 3 ans = $S(3)$ = (Survie instantanée entre 2 et 3 ans) x (survie instantanée entre 1 et 2 ans) x (survie instantanée entre 0 et 1 an)*

Instant	V	C	D	N= V- C/2	(N – D) / N	S(t)
0	-	-	-	-	-	1
3	210	0	0	210	1	1x1 = 1
6	210	10	40	210-5 = 205	(205-40)/205 = 0,805 →	0,805 x 1 = 0,805

Instant	V	C	D	N= V- C	(N – D) / N	S(t)
0	21	-	-	-	-	1
6	21	0	3	21	0,857	0,857
7	18	1	1	17	0,941	0,807

Courbe de la survie





Pour chaque intervalle de temps l'estimation de la survie est **représentée par un point**.	La courbe de survie se compose de **paliers successifs**, où les probabilités de survie sont constantes entre deux temps d'évènements consécutifs. *Ex : Le premier palier vaut 1 depuis l'origine jusqu'au délai de survenue du premier événement.*

D) Choix d'une valeur résumée

- **Médiane de survie :** Elle représente la durée t pour laquelle la probabilité de survie S(t) est de **50 %**. En pratique, la médiane est estimée par la plus petite durée pour laquelle la survie est inférieure à 50 %
- Quantiles de survie : Pour le p^{ième} quantile on estime la durée pour laquelle la probabilité de survie est de 100 – p
- Survie à date fixée : Estimation de la survie à un temps donné

E) Estimation de la fonction de survie

On cherche à estimer localement la fonction de survie S(ti) pour « i » fixé :

- **S(ti)** = Probabilité pour que le décès survienne après ti : P(T>ti)
- **S(ti-1)** = Probabilité pour que le décès survienne après ti-1 : P(T>ti-1)
- **S(ti/ti-1)** = Probabilité de survie = survie instantanée = $\frac{N-D}{N}$

$$\boxed{S(t_i) = S(t_i|t_{i-1}) \times S(t_{i-1})}$$

IV. Comparaison de deux fonctions de survie

On souhaite montrer qu'une action (intervention, traitement) ou une classification ont un lien avec la survie. Il s'agira de conduire une étude comparative et de mettre en œuvre un test d'hypothèses

A) Méthode du Log-rank

→ Principe : comparer, dans chaque groupe, le nombre **observé** et le nombre **attendu** d'événements si la survie était identique dans les deux groupes, sur l'ensemble de la période étudié

Attention aux biais ! Ne pas comparer la survie des patients **répondant** au traitement à la survie de ceux n'y répondant **pas** (dans ce cas les 2 groupes reçoivent bien le même traitement)
→ Il faut comparer la survie des patients **recevant** un traitement, à la survie de ceux qui n'en **reçoivent aucun**, ou un **différent**

Le test du log-rank :

- Test du chi^2 à un degré de liberté
- H0 : Les fonctions de survie sont les mêmes dans les 2 groupes S(t/A) = S(t/B)
- H1 : Les fonctions de survies sont différentes
- Tests généralisables à plusieurs groupes : permet de tester si globalement la survie est différente entre les groupes

Je sais que ce cours paraît assez indigeste mais les questions posées ne seront pas pointilleuses, dures ou quoi que ce soit ! Le professeur préfère rester dans les « généralités » et ses cours de la fin sont des points faciles, donc à ne pas délaisser !
Je vous souhaite un gros gros gros bon courage à tous pour ce dernier mois, n'abandonnez pas et donnez le meilleur de vous-même quel que soit votre situation !

Petit mot pour mes préférés : Sonia, laure-Anne, Anthony, Maud, Alexiane, Marie, Manon, Sophie, Emilie, Vincent, si l'un d'entre vous ne lis pas cette fiche je l'étrangle ! Non plus sérieusement, je crois en vous ♡

« Au moment où tu abandonnes tu laisses quelqu'un d'autre gagner »