

MINI FICHE RÉCAP - ANALYSE DE SURVIE

Analyse actuarielle

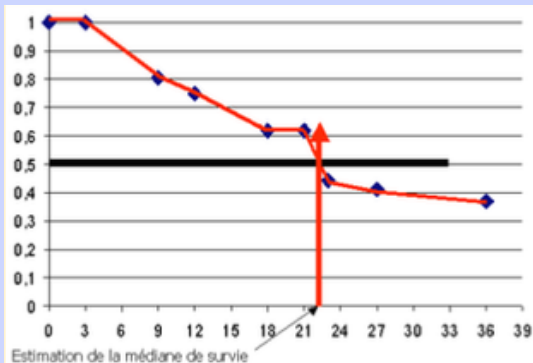
Moins utilisée

Grand nombre de sujets : $n > 200$

Suppose que les sujets soient exposés au risque d'évènement sur la moitié de l'intervalle

Fonction de survie calculée sur des intervalles de temps fixés à priori

$$N = V - C/2$$



Méthode Kaplan-Meier

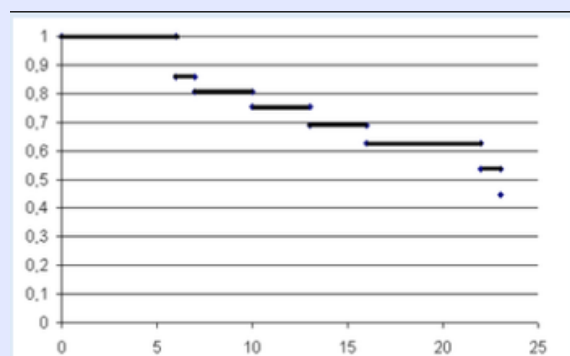
Plus utilisée

Échantillon de taille réduite

/

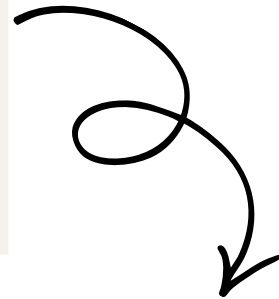
Les intervalles sont définis par les instants où les évènements sont observés

$$N = V - C$$



Pour chaque intervalle de temps entre 2 dates, on définit :

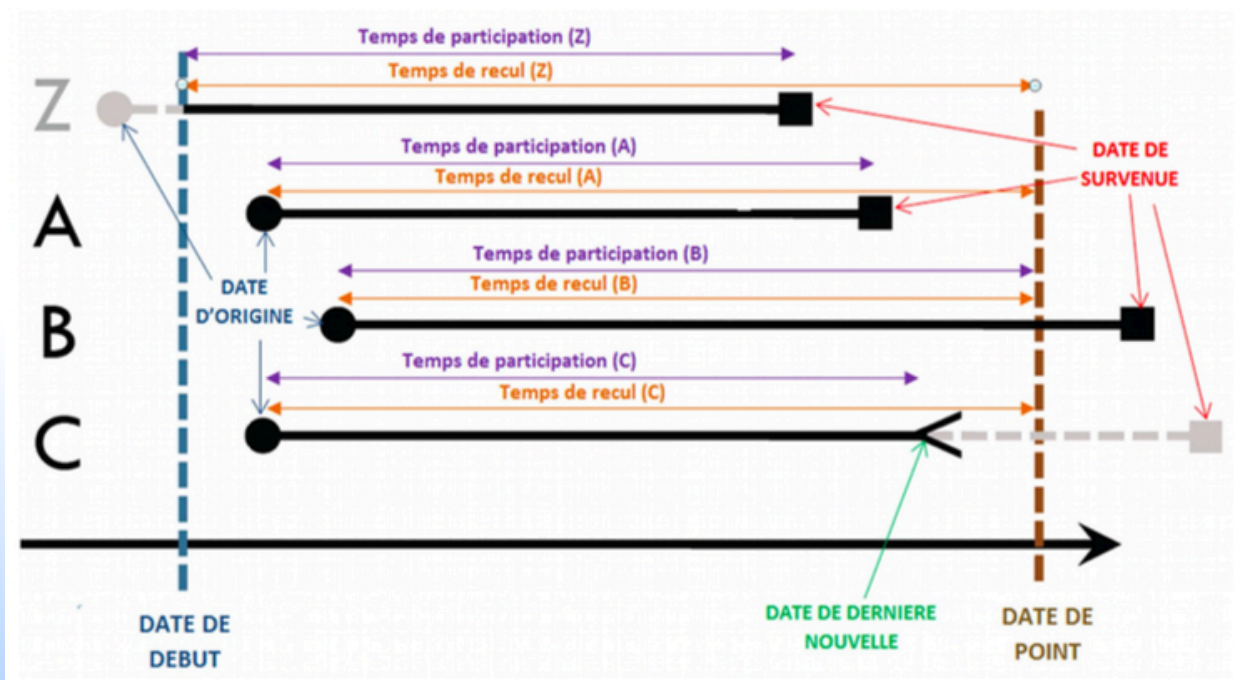
- **V** : nombre de sujets vivants au début de l'intervalle
- **D** : nombre de sujets décédés dans l'intervalle
- **C** : censurés car perdus de vue
- **N** : sujets exposés au risque d'évènements sur l'intervalle



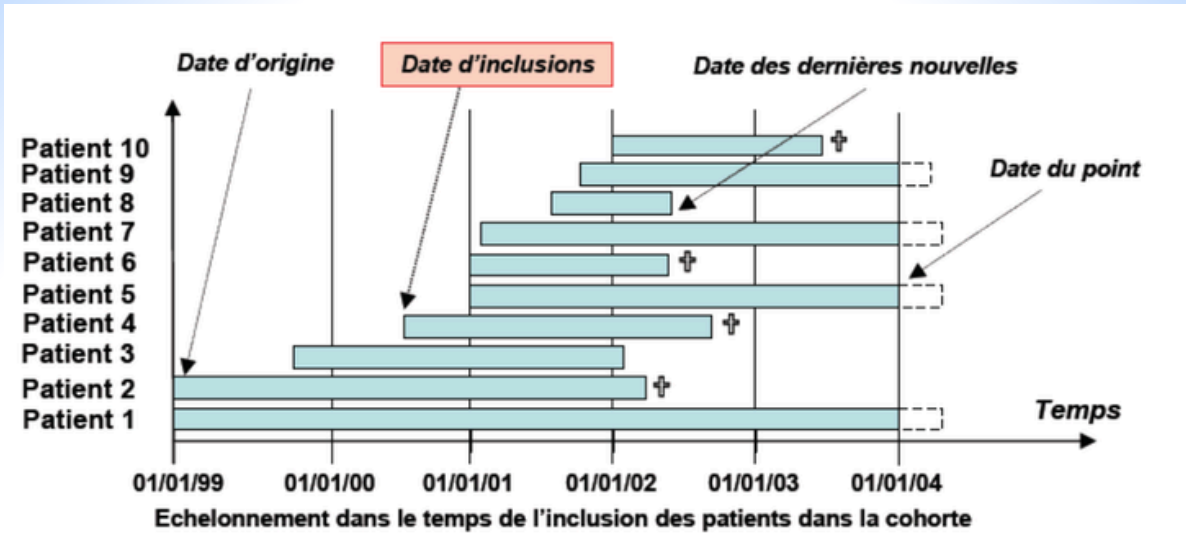
On peut avoir un tableau de données comme ceci

Instants	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
10	16	1	1	15	0,933	0,753
13	14	2	1	12	0,917	0,690
16	11	0	1	11	0,909	0,627
22	10	3	1	7	0,857	0,537
23	6	0	1	6	0,833	0,448

La probabilité d'évènement durant l'intervalle : D/N
 La survie sur l'intervalle → survie instantanée : $(N-D)/N$



À propos des différentes dates :



Les **dates à connaître** : date de début d'étude, date de fin d'étude (= date du point), date d'origine (=date d'inclusion) et date des dernières nouvelles.

Autres notions : Temps de recul et temps de participation

La **date de début d'étude** ici c'est le 1^{er} janvier 1999, soit quand on commence à observer tes premiers patients à savoir les patients 1 et 2.

La **date de point** ici c'est le 1^{er} janvier 2004, c'est la date où on fait le bilan. C'est aussi là qu'on va pouvoir déterminer toutes les durées de survie des individus qui seront censurés. En effet, après cette date les informations sur les patients ne nous intéressent pas (dans le cadre de l'analyse). Ainsi les données de survies des patients 1, 5, 7 et 9 sont censurées car l'événement d'intérêt (ici le décès) n'est pas observé, ce sont les exclus-vivants. En effet ces derniers sont toujours vivants à la date de point. Tout ce qui se passe après cette date n'est plus de notre ressort (en tout cas dans l'analyse prévue).

La **date d'origine** correspond où commence la surveillance des sujets, le plus souvent elle correspond à la date d'inclusion des sujets. Tous ne sont pas forcément inclus au même moment. Par exemple, la date d'origine du patient 2 est le 1^{er} janvier 1999 (idem date début d'étude) alors que la date d'origine pour le patient 6 est le 1^{er} janvier 2001.

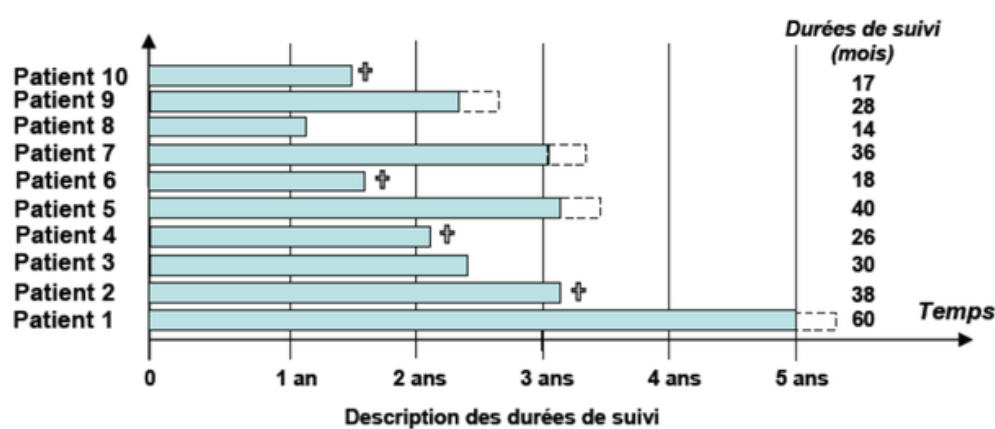
Cas particulier vu dans le cours : patient Z, où la date d'origine est antérieure à l'inclusion dans l'étude. Je vais reprendre le même exemple : on décide d'inclure des sujets dans notre étude quand on observe qu'ils sont hypertendus. Pour certains cela va être récent donc leur date d'inclusion/d'origine va être la même. Pour d'autres (cas Z), ils sont hypertendus depuis bien longtemps, avant même le début de l'étude, leur date d'origine va correspondre à l'apparition de leur hypertension.

La **date des dernières nouvelles** correspond à la date la plus récente où on a pu recueillir des informations sur nos sujets. Par exemple, pour le patient 1 la date des dernières nouvelles correspond à la date de point (1^{er} janvier 2004 – c’est ce qu’on appelle un exclu vivant, idem patients 5, 7 et 9), pour le patient 2 la date de ses dernières nouvelles correspond à son décès (événement d’intérêt – c’est bueno on ne censure pas). Pour le patient 3 la date des dernières nouvelles c’est un peu après le 1^{er} janvier 2002, lui c’est un perdu de vue car l’événement d’intérêt n’a pas été observé et qu’on ne l’a pas suivi jusqu’à la date de point (de même pour le patient 8).

Finalement on censure les durées de survies des sujets chez qui l’événement d’intérêt n’a pas été observé : à savoir les patients 1, 3, 5, 7, 8 et 9.

Le **temps de recul** va correspondre au délai entre la date d’origine et la date de point. En cela, les patients 1 et 2 ont le même temps de recul car la date de point est la même pour tous les patients et que leur date d’origine est la même. Aussi on peut dire que leur temps de recul est supérieur aux autres patients (ils sont inclus dans l’étude avant les autres patients). En gros, c’est un peu comme si on se demandait le jour de la date de point : « depuis combien de temps je t’ai inclus ? »

Le **temps de participation** c’est la durée de surveillance pour chaque sujet. C’est ce qui nous vient plus intuitivement par rapport au temps de recul. C’est ce que tu peux observer grâce aux barres bleues. C’est comme son nom l’indique le temps que tu passes à participer à l’étude, c’est la durée de suivi de ton patient. On peut voir sur le graphique que le patient 8 a un plus petit temps de participation que les autres. Quand tu utilises un calendrier statisticien (aux durées relatives) c’est encore plus visible.



Sur l’axe des abscisses c’est le temps de participation en années.
À droite le temps de participation (=durée de suivi) en mois.