



Cours n° 3 : Analyse de la survie

By Minipousse

- Le tutorat est gratuit. Toute reproduction ou vente sont interdites.

INTRODUCTION :

- Objectifs :
 - estimer et expliquer la durée de survie en fonction de certains facteurs
 - comparer la survie entre plusieurs groupes de sujets

L'évènement d'intérêt est souvent le décès mais peut être :

- ▶ -la survenue d'une complication
- ▶ -la rechute d'une maladie
- ▶ -la disparition de symptômes

Une étude de survie est une étude :

- ▶ Longitudinale
- ▶ Prospective
- ▶ De cohorte

I/ Quelques définitions...importantes !!

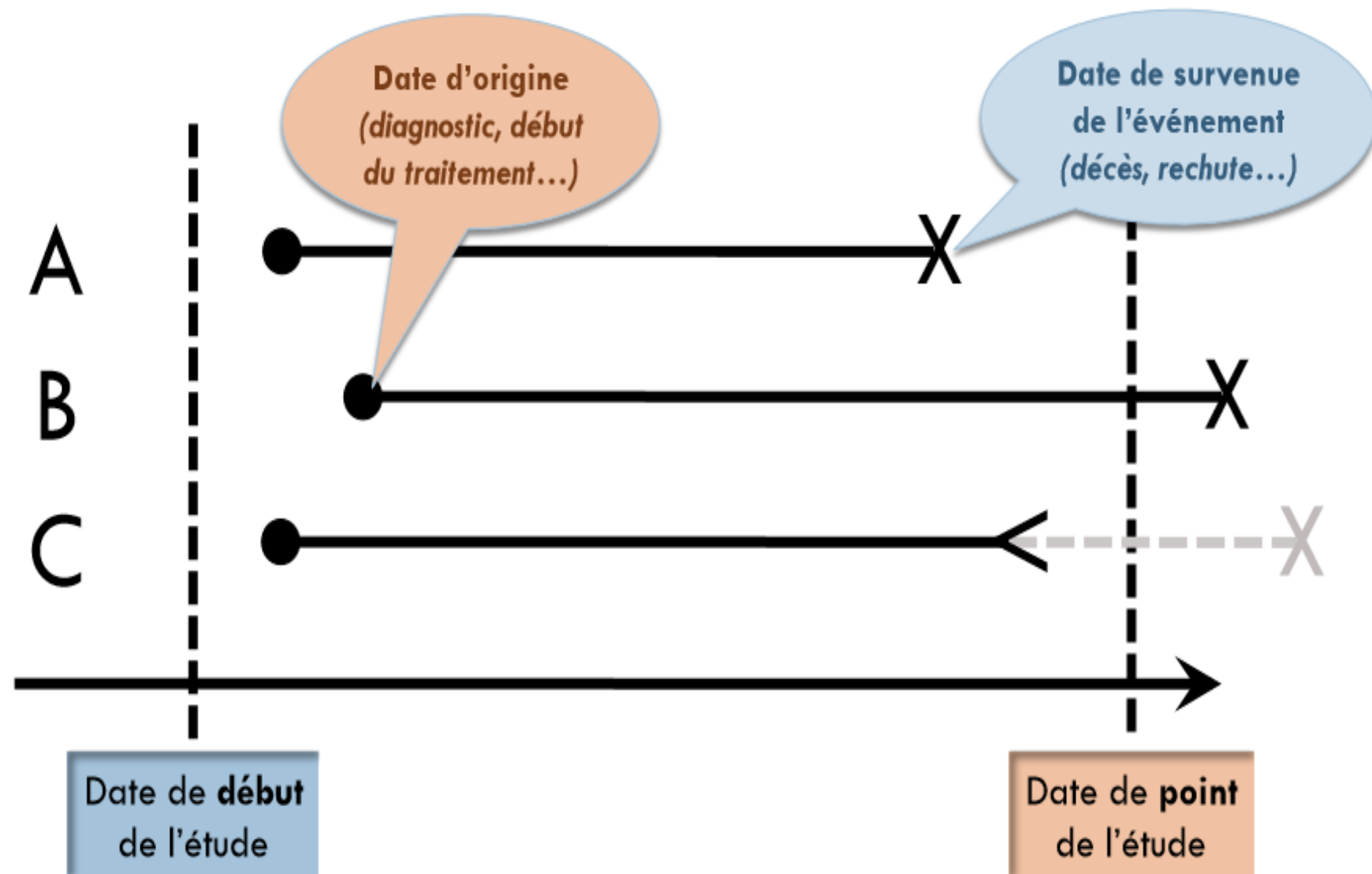
- ▶ **Cohorte** : Ensemble de sujets inclus dans une étude au même moment, et suivis dans des conditions standardisées pendant une durée prédéfinie.

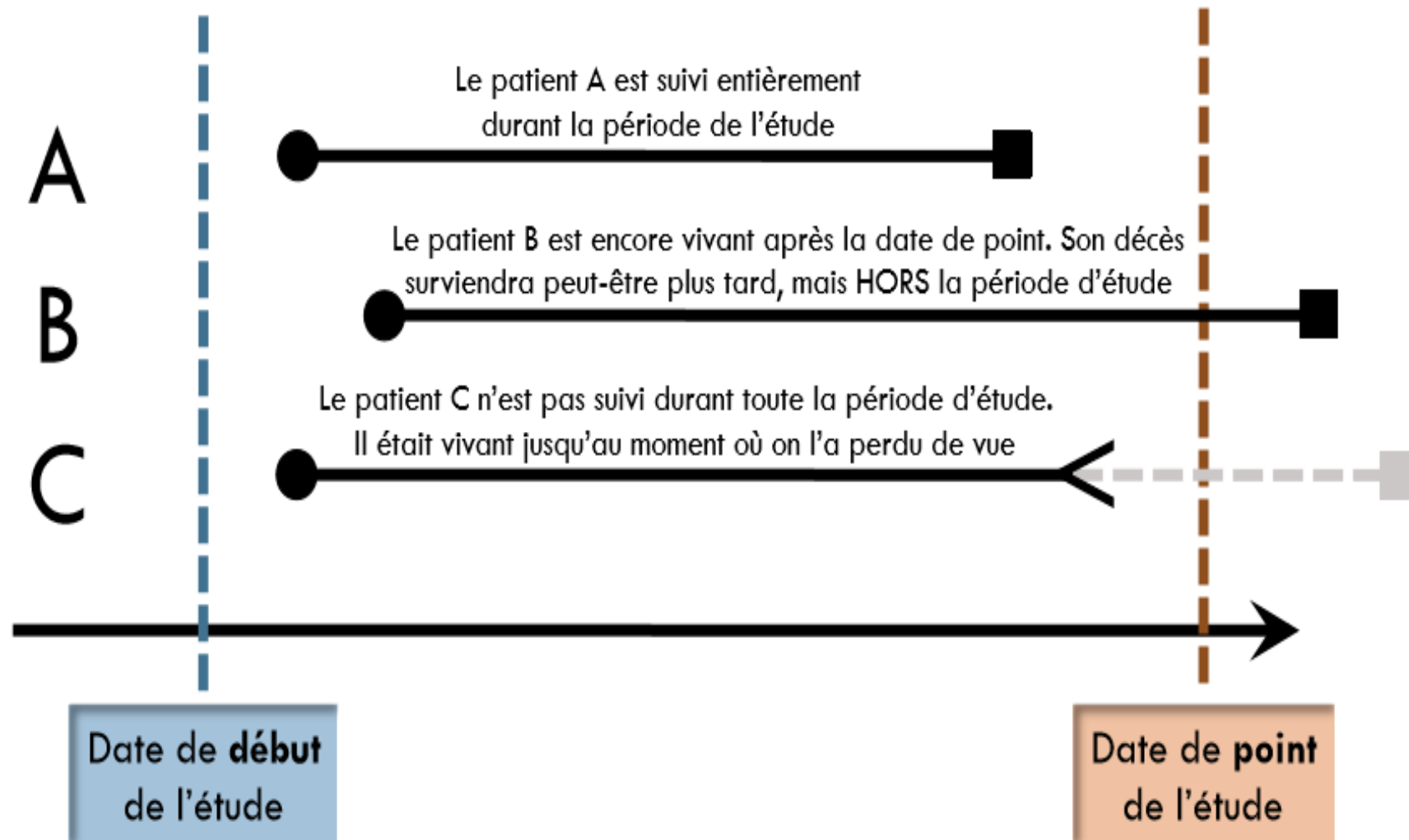
- ▶ **Cohorte « incipiente »** : la cohorte des patients qui rentrent dans l'étude doit inclure des sujets observés au début de leur affection au même stade de leur maladie (« cas incidents »).

- ▶ **Date d'origine** : C'est la date correspondant au point de départ de la surveillance.
- ▶ **Date de point** : C'est la date choisie pour faire le bilan.

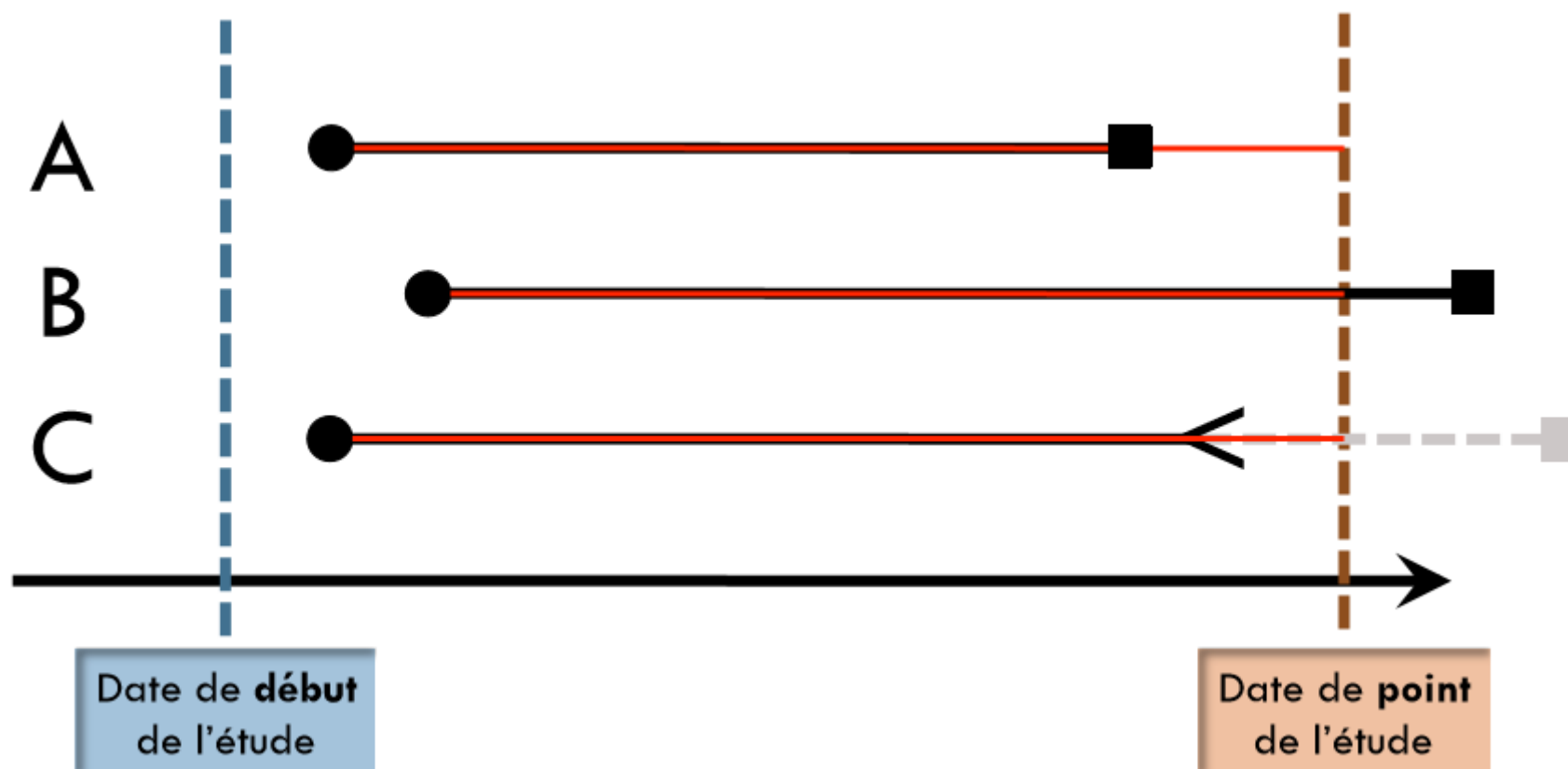
- ▶ **Censure** : Une durée de survie d'un individu est dite censurée lorsque l'événement d'intérêt n'a pas été observé.

- ▶ **Date des dernières nouvelles:** 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.
- ▶ **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.

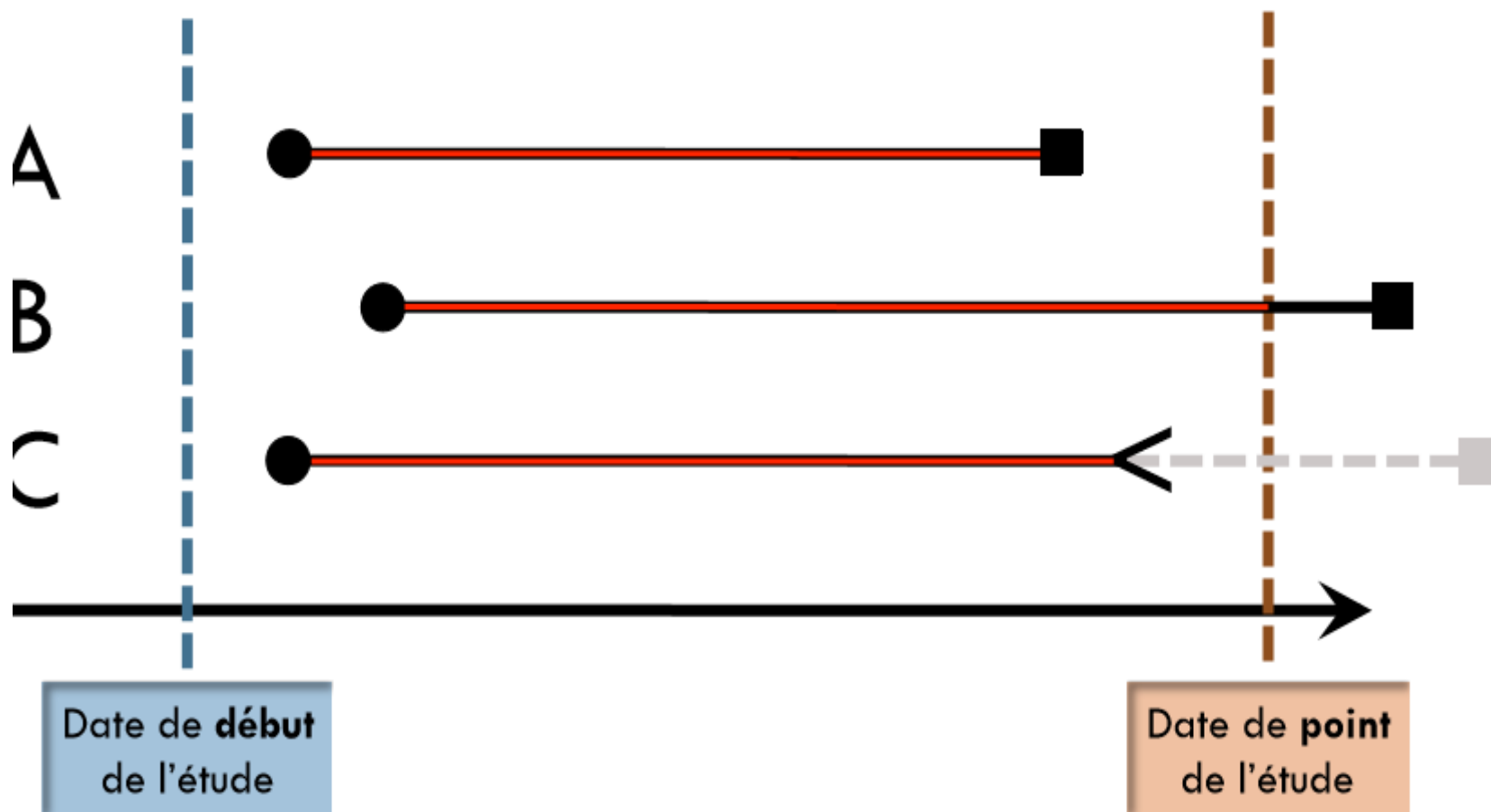




- **Temps de recul** : Délai entre la date d'origine et la date de point, c'est-à-dire le délai maximum potentiel de suivi pour un sujet.



- **Temps de participation** : Durée de surveillance pour chaque sujet utilisée dans l'estimation de la survie.



II/ La fonction de survie :

► Loi exponentielle :

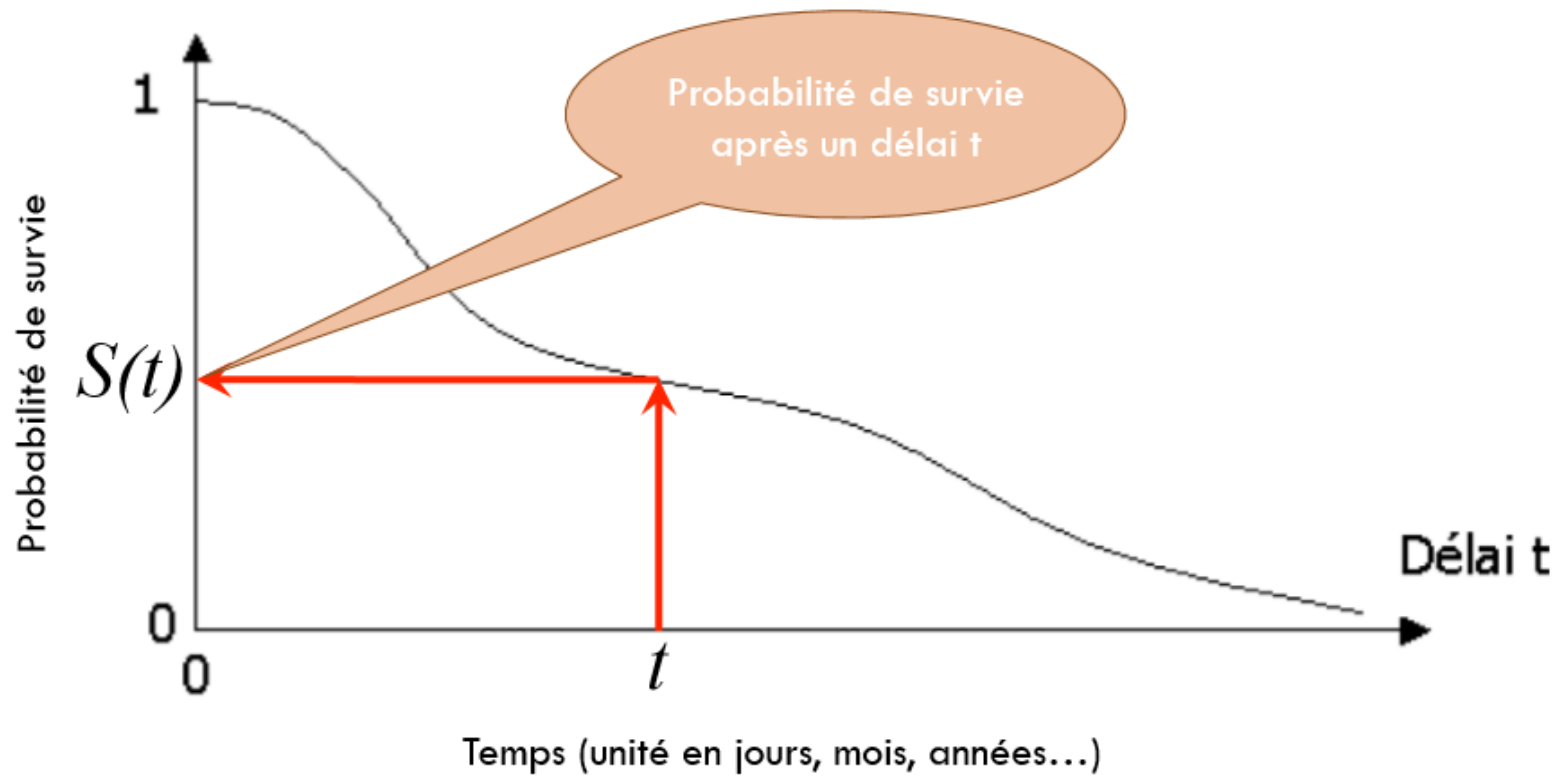
Fonction de répartition :

$$F(t) = P(X \leq t) = \int_0^t \lambda e^{-\lambda x} dx = 1 - e^{-\lambda t}$$

- ▶ La quantité $1-F(t)$ représente donc la quantité d'équipements qui fonctionnent pendant une durée égale au moins à t .
- ▶ La fonction de survie :

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

Courbe de survie :



- Le tutorat est gratuit. Toute reproduction ou vente sont interdites.

Probabilités:

- ▶ La probabilité pour que le décès survienne après un délai t_1 et avant un délai t_2 ($t_2 > t_1$) :

$$Pr(T \in]t_1; t_2]) = F(t_2) - F(t_1) = S(t_1) - S(t_2)$$

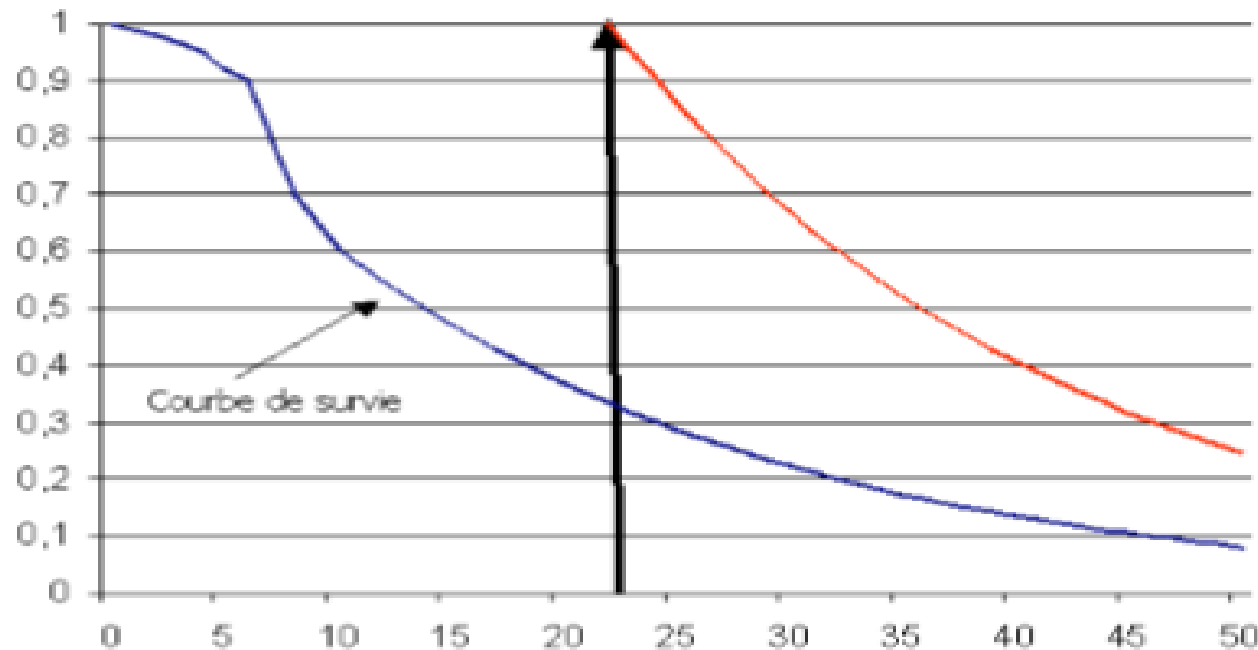
► La probabilité de survivre encore après un délai « t » sachant que l'on est survivant après un délai « τ » ($\tau < t$) se note $S(t/\tau)$.

Ainsi :

$$S(t/\tau) = \frac{S(t)}{S(\tau)}$$

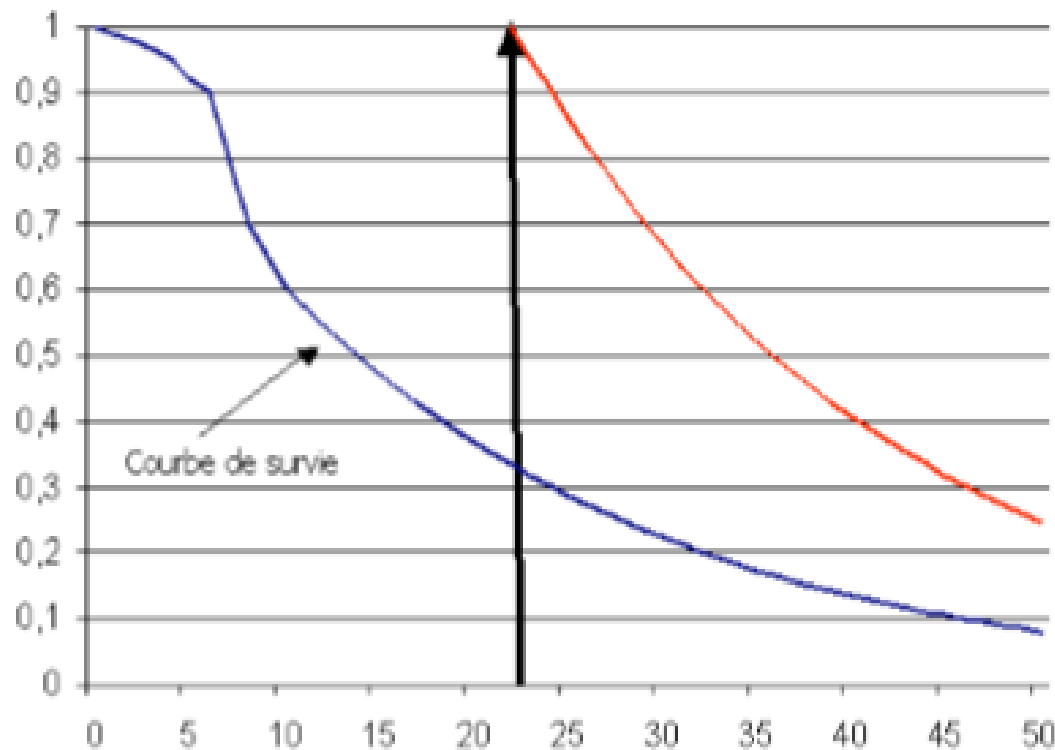
Exemple :

- ▶ Quelle est la probabilité de survivre après un délai de $t=33$ ans sachant que l'on est vivant à $t=23$ ans ?



► $S(23\text{ans}) = 0,33$ et $S(33\text{ans}) = 0,2$.

Ainsi, $S(33\text{ans})/S(23\text{ans}) = 0,2/0,33 = 0,6$



III/ Estimation de la survie :

On utilise deux méthodes non paramétriques.

- ▶ **L'analyse actuarielle** si les échantillons sont grands (+ de 200 sujets)
- ▶ **La méthode de Kaplan-Meier** si les échantillons sont plus petits (- de 200 sujets)

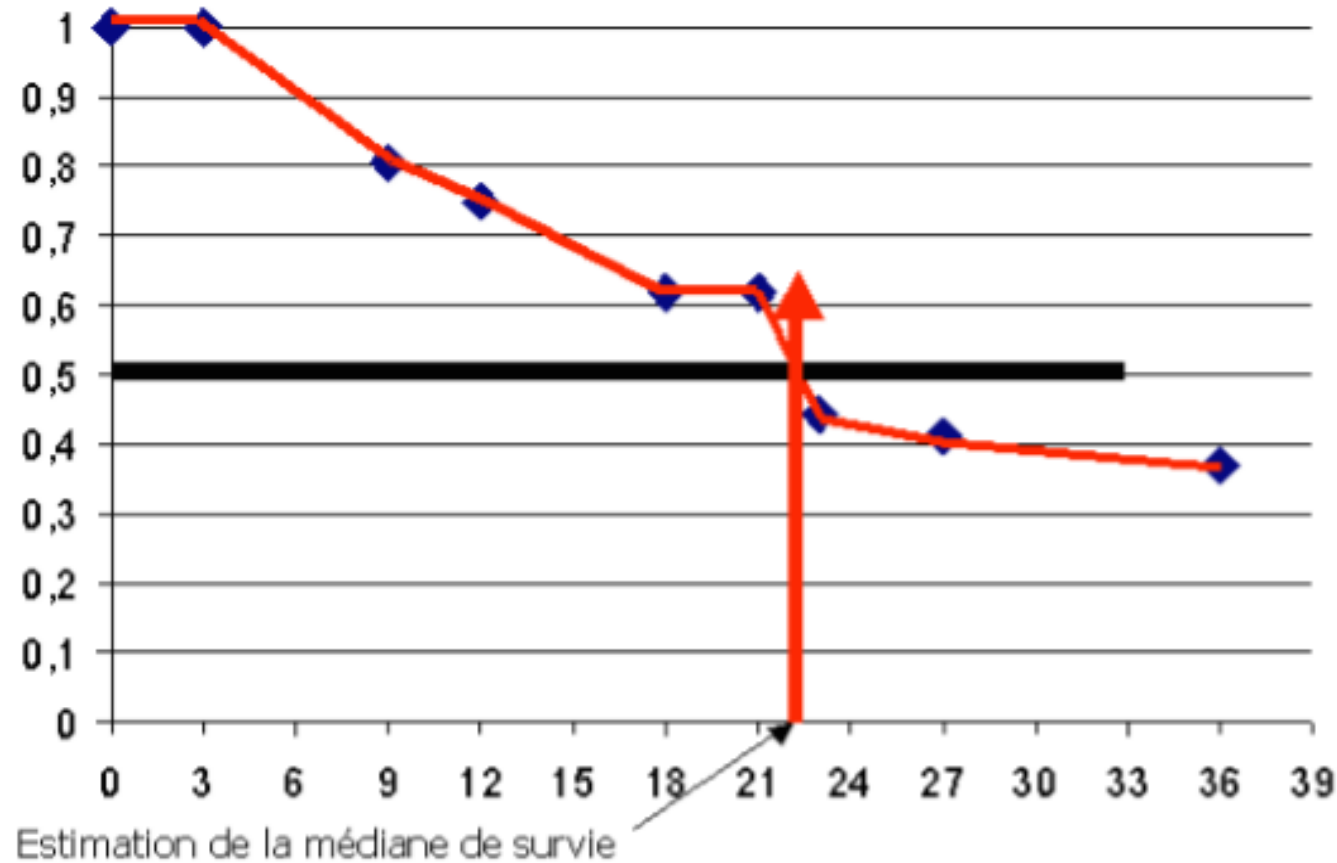
L'analyse actuarielle :

- ▶ Les intervalles de temps sont fixés à priori (mois, trimestre, semestre, année...).
- ▶ La survie sur un intervalle est appelée survie instantanée.

Instants	V	C	D	$N = V - C/2$	$(N - D) / N$	$S(t)$
0	-	-	-	-	-	1
3	210	0	0	210	1	$1 \times 1 = 1$
6	210	10	40	$210 - 5 = 205$	$(205-40)/205 = 0,805$	$0,805 \times 1 = 0,805$
9	160	30	10	$160 - 15 = 145$	$(145-10)/145 = 0,931$	$0,931 \times 0,805 = 0,749$
12	120	10	20	$120 - 5 = 115$	$(115-20)/115 = 0,826$	$0,826 \times 0,749 = 0,619$
15	90	20	0	$90 - 10 = 80$	1	$1 \times 0,619 = 0,619$
18	70	0	20	70	$(70-20)/70 = 0,714$	$0,714 \times 0,619 = 0,442$
21	50	18	3	$50 - 9 = 41$	$(41-3)/41 = 0,927$	$0,927 \times 0,442 = 0,410$
24	29	8	2	$29 - 4 = 25$	$(25-2)/25 = 0,920$	$0,920 \times 0,410 = 0,377$

V : nombre de sujets vivants au début de l'intervalle
 C : nombre de sujets vivants censurés dans l'intervalle
 D : nombre de sujets décédés dans l'intervalle
 N : nombre de sujets exposés au risque de décès

Pour chaque intervalle de temps, on représente l'estimation de la survie $S(t)$ par un point.



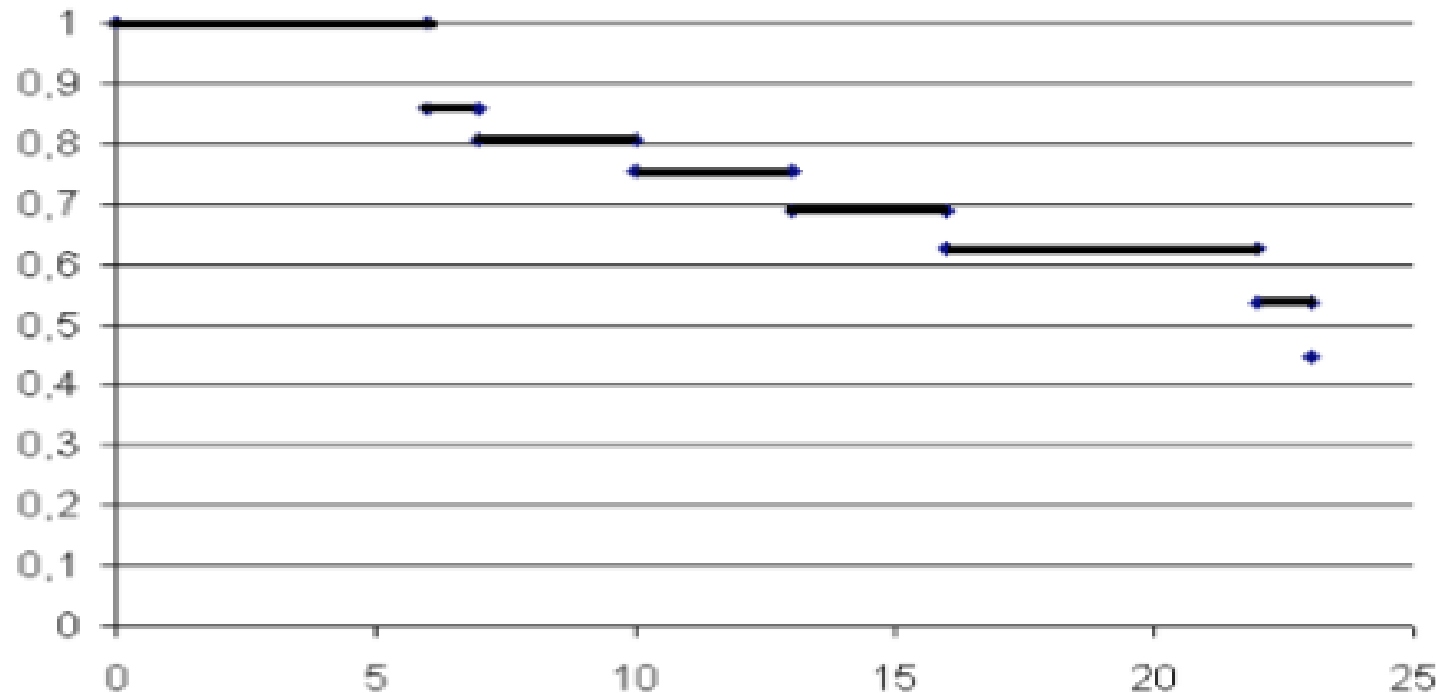
La méthode de Kaplan-Meier:

- ▶ les intervalles ne sont pas ici fixés a priori, mais sont définis par les instants auxquels les événements sont observés.

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

V : nombre de sujets vivants au début de l'intervalle
 C : nombre de sujets vivants censurés dans l'intervalle
 D : nombre de sujets décédés dans l'intervalle
 N : nombre de sujets exposés au risque de décès

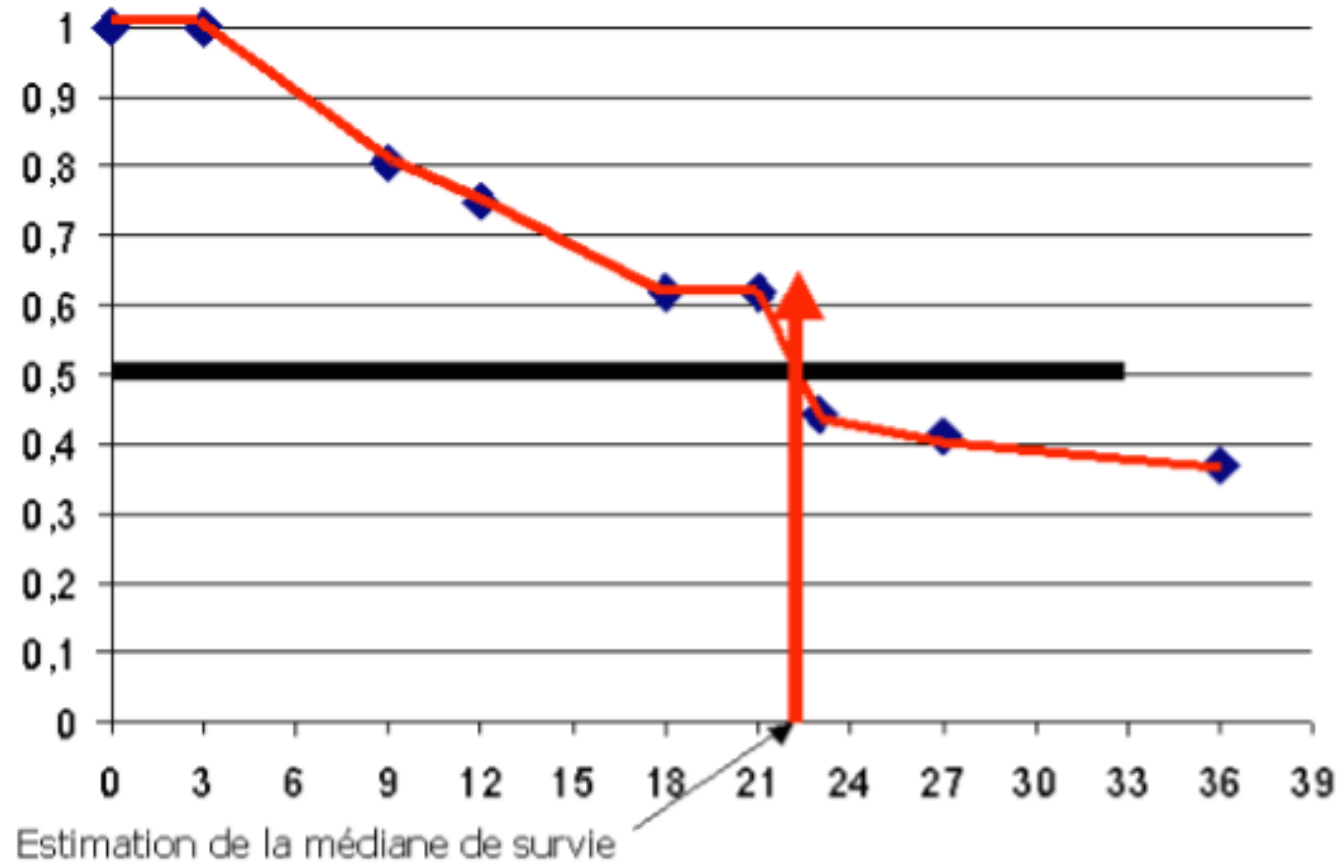
La courbe de survie se compose de paliers successifs, où les probabilités de survie sont constantes entre deux temps d'événements consécutifs. Le premier palier vaut 1 depuis l'origine jusqu'au délai de survenue du premier événement.



Le choix d'une valeur résumée :

- ▶ **La médiane de survie** est plus significative que la moyenne 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 %.

Pour chaque intervalle de temps, on représente l'estimation de la survie $S(t)$ par un point.



Instants	V	C	D	$N = V - C/2$	$(N - D) / N$	S(t)
0	-	-	-	-	-	1
3	210	0	0	210	1	$1 \times 1 = 1$
6	210	10	40	$210 - 5 = 205$	$(205-40)/205 = 0,805$	$0,805 \times 1 = 0,805$
9	160	30	10	$160 - 15 = 145$	$(145-10)/145 = 0,931$	$0,931 \times 0,805 = 0,749$
12	120	10	20	$120 - 5 = 115$	$(115-20)/115 = 0,826$	$0,826 \times 0,749 = 0,619$
15	90	20	0	$90 - 10 = 80$	1	$1 \times 0,619 = 0,619$
18	70	0	20	70	$(70-20)/70 = 0,714$	$0,714 \times 0,619 = 0,442$
21	50	18	3	$50 - 9 = 41$	$(41-3)/41 = 0,927$	$0,927 \times 0,442 = 0,410$
24	29	8	2	$29 - 4 = 25$	$(25-2)/25 = 0,920$	$0,920 \times 0,410 = 0,377$

V : nombre de sujets vivants au début de l'intervalle
 C : nombre de sujets vivants censurés dans l'intervalle
 D : nombre de sujets décédés dans l'intervalle
 N : nombre de sujets exposés au risque de décès

QCM :



Le tutorat est gratuit. Toute reproduction ou vente sont interdites

1) Une étude de survie doit être :

- ▶ A) Horizontale
- ▶ B) De cohorte
- ▶ C) Rétrospective
- ▶ D) Longitudinale
- ▶ E) Les réponses A,B,C et D sont fausses.

Réponses B et D !!!!

C : PROSPECTIVE

simple question de cours ;)

2) Dans une étude de survie, le temps de participation d'un individu perdu de vue :

- ▶ A) Est le délai entre la date d'origine et la date de survenue de l'évènement
- ▶ B) Est le délai entre la date d'origine et la date de point
- ▶ C) Est le délai entre la date d'origine et la date des dernières nouvelles
- ▶ D) Ne peut être déterminé
- ▶ E) Les réponses A,B,C et D sont fausses.

Réponse C !!!!!!!

Le tutorat est gratuit. Toute reproduction ou vente sont interdites

3) On s'intéresse à la survie d'un échantillon de 100 patients atteints d'une maladie infectieuse pouvant être mortelle. Les résultats de l'étude sont consignés dans le tableau suivant au terme d'un suivi qui a duré un an et trois mois.

Instants	V	C	D	$N = V - C$	D / N	$(N - D) / N$	$S(t)$
0	100	-	-	-	-	-	1
3 mois	100	0	19	100	0,190	0,810	0,810
6 mois	81	2	12	79	0,152	0,848	0,687
9 mois	67	1	7	66	0,106	0,894	0,614
12 mois	59	1	3	58	0,052	0,948	0,582
15 mois	55	0	1	55	0,018	0,982	0,571

V : nombre de sujets vivants au début de l'intervalle

C : nombre de sujets vivants censurés dans l'intervalle

D : nombre de sujets décédés dans l'intervalle

N : nombre de sujets exposés au risque de décès dans l'intervalle

Instants	V	C	D	N = V-C	D / N	(N-D) / N	S(t)
0	100	-	-	-	-	-	1
3 mois	100	0	19	100	0,190	0,810	0,810
6 mois	81	2	12	79	0,152	0,848	0,687
9 mois	67	1	7	66	0,106	0,894	0,614
12 mois	59	1	3	58	0,052	0,948	0,582
15 mois	55	0	1	55	0,018	0,982	0,571

V : nombre de sujets vivants au début de l'intervalle
C : nombre de sujets vivants censurés dans l'intervalle
D : nombre de sujets décédés dans l'intervalle
N : nombre de sujets exposés au risque de décès dans l'intervalle

- A)Le test utilisé est paramétrique, il s'agit de la méthode de Kaplan-Meier
- B)Le test utilisé est non paramétrique, il s'agit de l'analyse actuarielle
- C)La survie instantanée à 15 mois est de 98,2 %
- D)La probabilité d'être en vie à 12 mois est de 94,8 %
- E)Les réponses A,B,C et D sont fausses.

Réponse C !!!!!

- ▶ A et B fausses car les tests utilisés sont non paramétriques. De plus c'est un petit effectif donc il s'agit de la méthode de Kaplan-Meier.
- ▶ D fausse car la probabilité énoncée correspond à la survie instantanée et non la fonction de survie !

Ouf c'est fini !!!!!

Merci pour votre
écoute et bon
courage !!!!!

