Estimation d'un

Taux de Survie

Pr Roch Giorgi

roch.giorgi@univ-amu.fr

SESSTIM, Faculté de Médecine, Aix-Marseille Université, Marseille, France
http://sesstim.univ-amu.fr/

Objectifs (1)

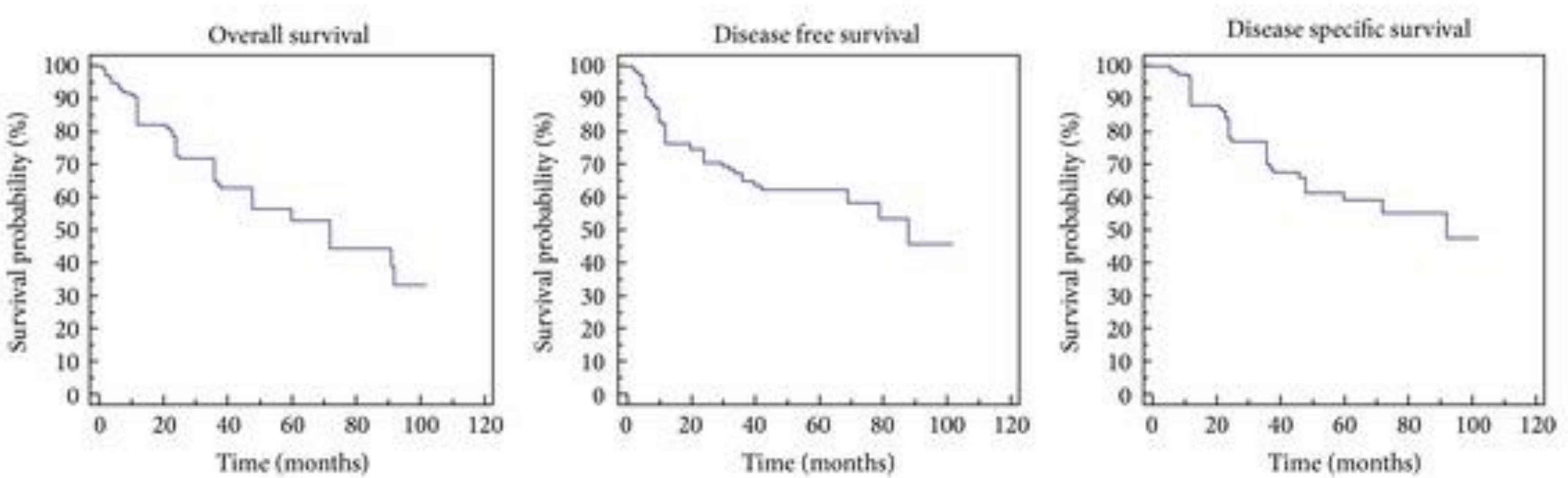
- Médical
  - ✓ Quantifier la probabilité de décès, de rechute, …
  - ✓ Évaluer les effets de facteurs pronostiques
  - ✓ Comparer l’efficacité de traitements

Objectifs (2)

- Statistique
  - ✓ Mesure du temps écoulé entre deux évènements
  - ✓ Prise en compte de données incomplètes : données censurées et tronquées
  - ✓ Distribution souvent dissymétrique (non Gaussienne)
  - ✓ Utilisation de méthodes non paramétriques

Principales Evolutions d'Intérêts

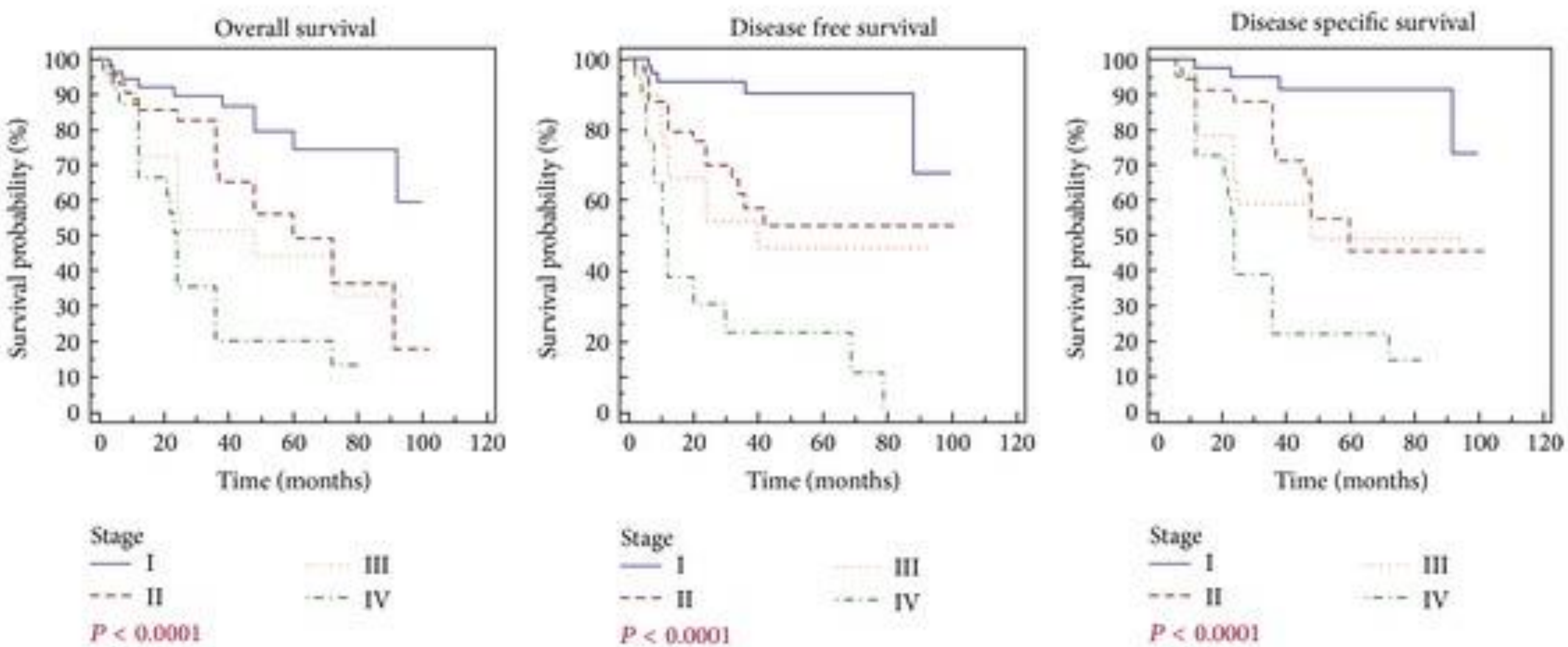
- Survie globale
- Survie spécifique
- Survie sans rechute



Source : Lorenzon L, Mercantini P, Ferri M et al. Profiling the prognosis of gastric cancer patients: is it worth correlating the survival with the clinical/pathological and molecular features of gastric cancers? *ScientificWorldJournal* 2013;2013:196541.

Principales Evolutions d'Intérêts

- Survie globale
- Survie spécifique
- Survie sans rechute



Source : Lorenzon L, Mercantini P, Ferri M et al. Profiling the prognosis of gastric cancer patients: is it worth correlating the survival with the clinical/pathological and molecular features of gastric cancers? *ScientificWorldJournal* 2013;2013:196541.

Temps et Évènement

- Temps : notion d'origine et de fin (évènement)
- Évènement : survie n'implique pas décès

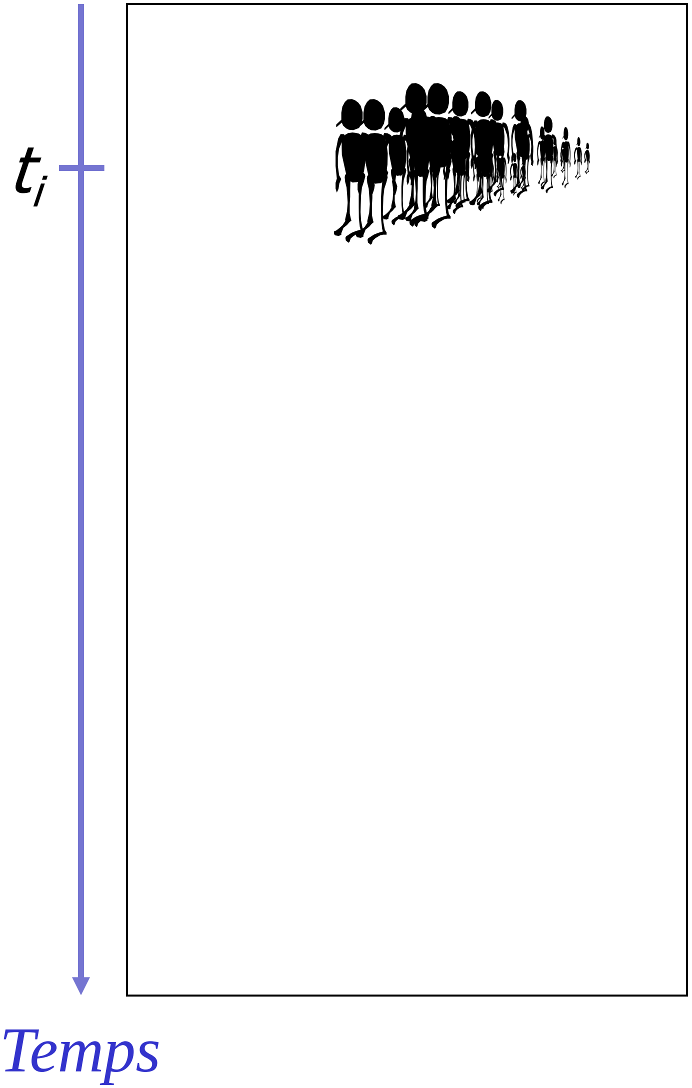
Origine

- ✓ Diagnostic
- ✓ Rémission
- ✓ Randomisation, Traitement
- ✓ Naissance
- ✓ Infection HIV

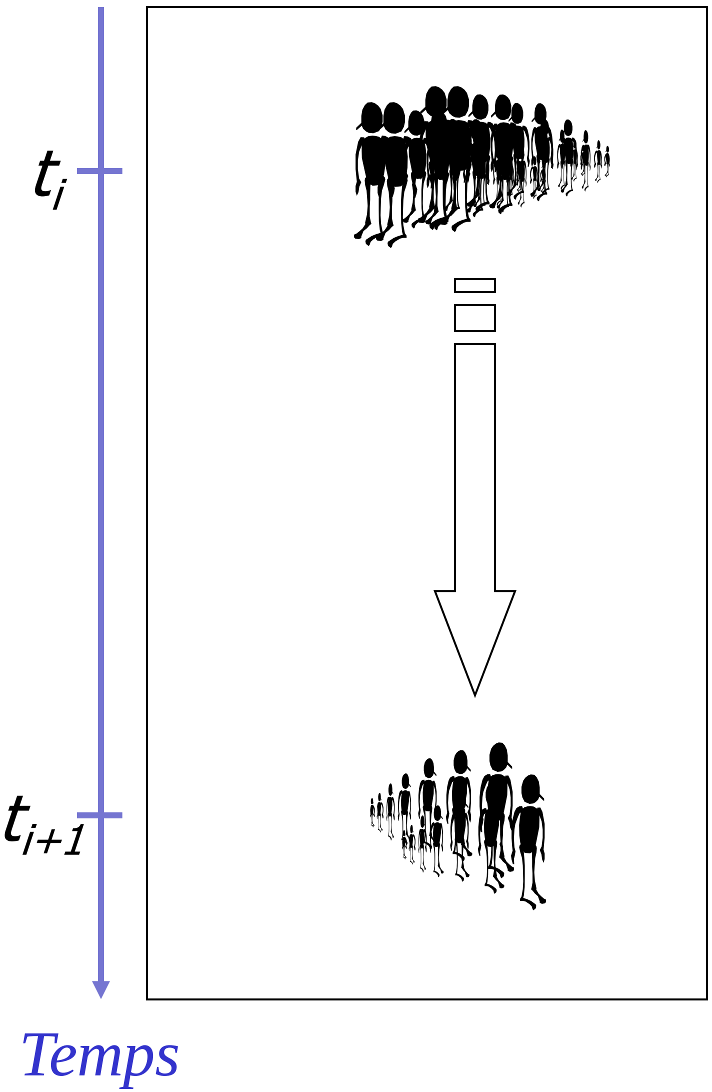
Évènement

- ✓ Décès
- ✓ Rechute
- ✓ Guérison, décès
- ✓ Maladie, décès
- ✓ SIDA

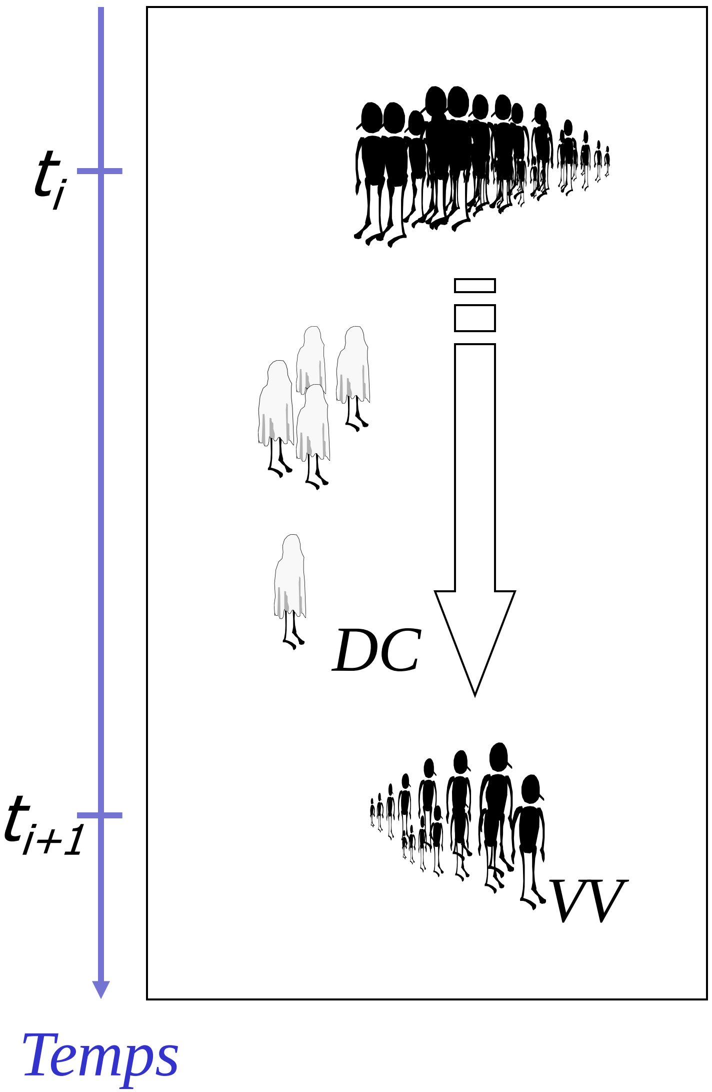
Observations Complètes



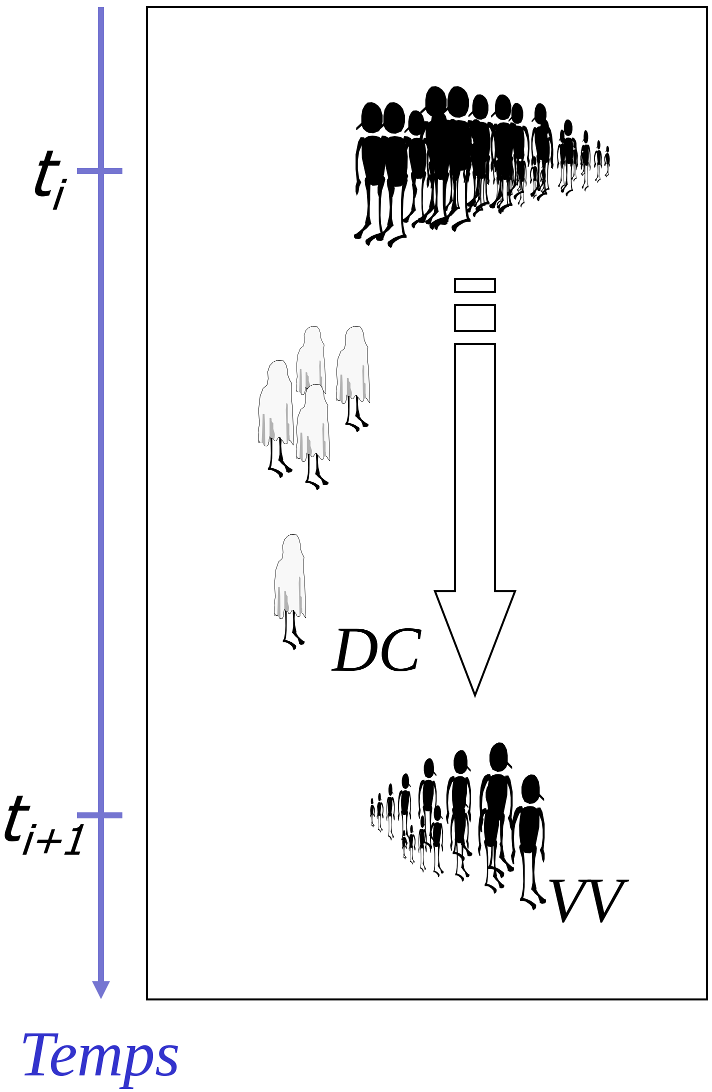
Observations Complètes



Observations Complètes



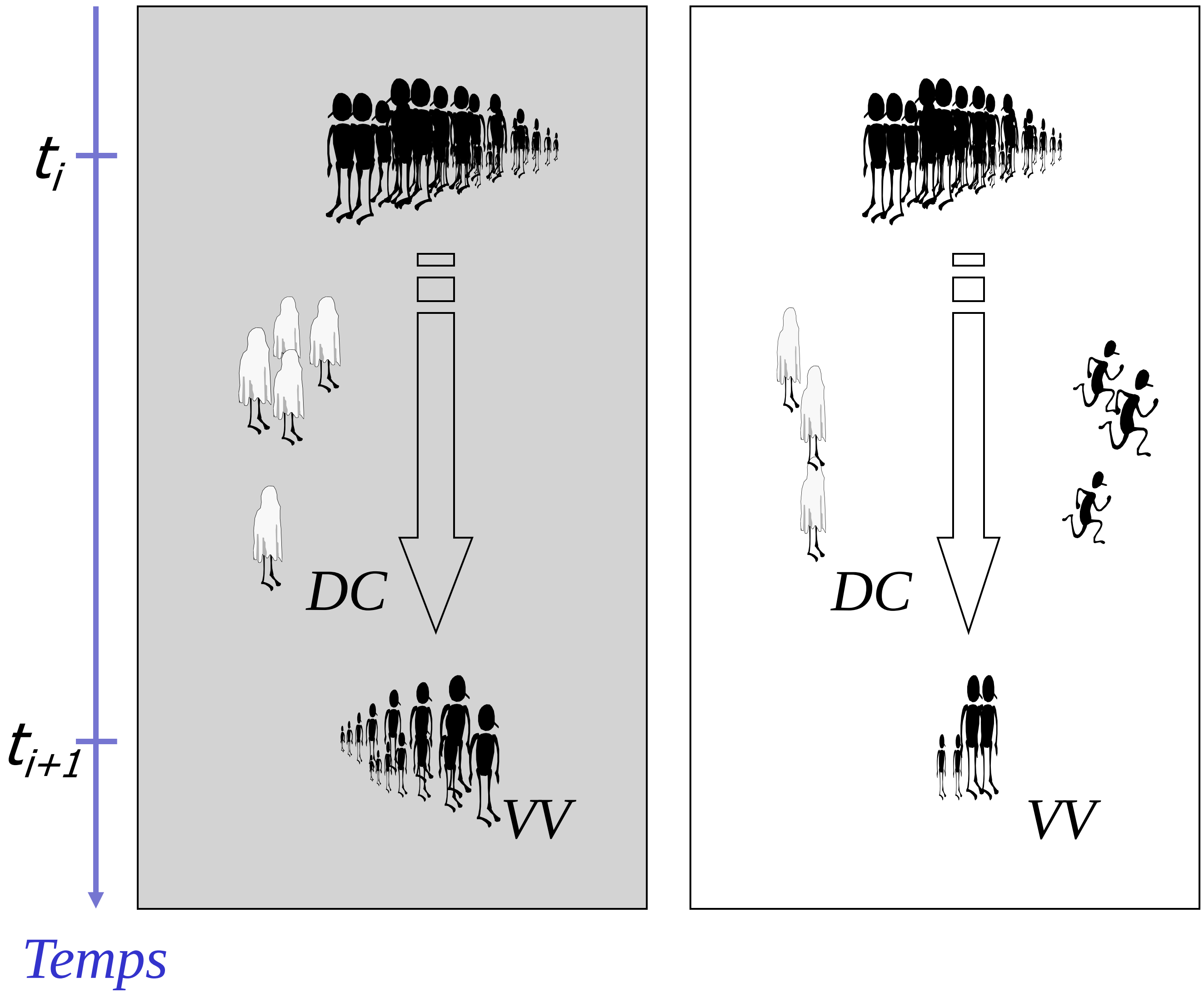
Observations Complètes



Proba. de Survie

$$= \frac{VV}{VV + DC}$$

Observations Incomplètes – Censures



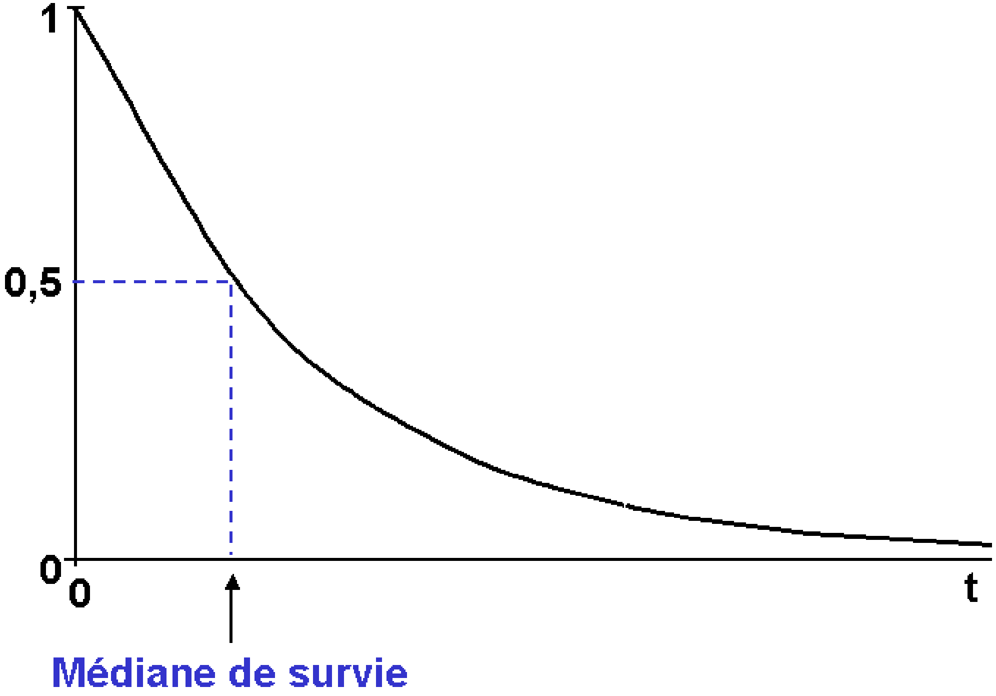
Proba. de Survie

$$\neq$$

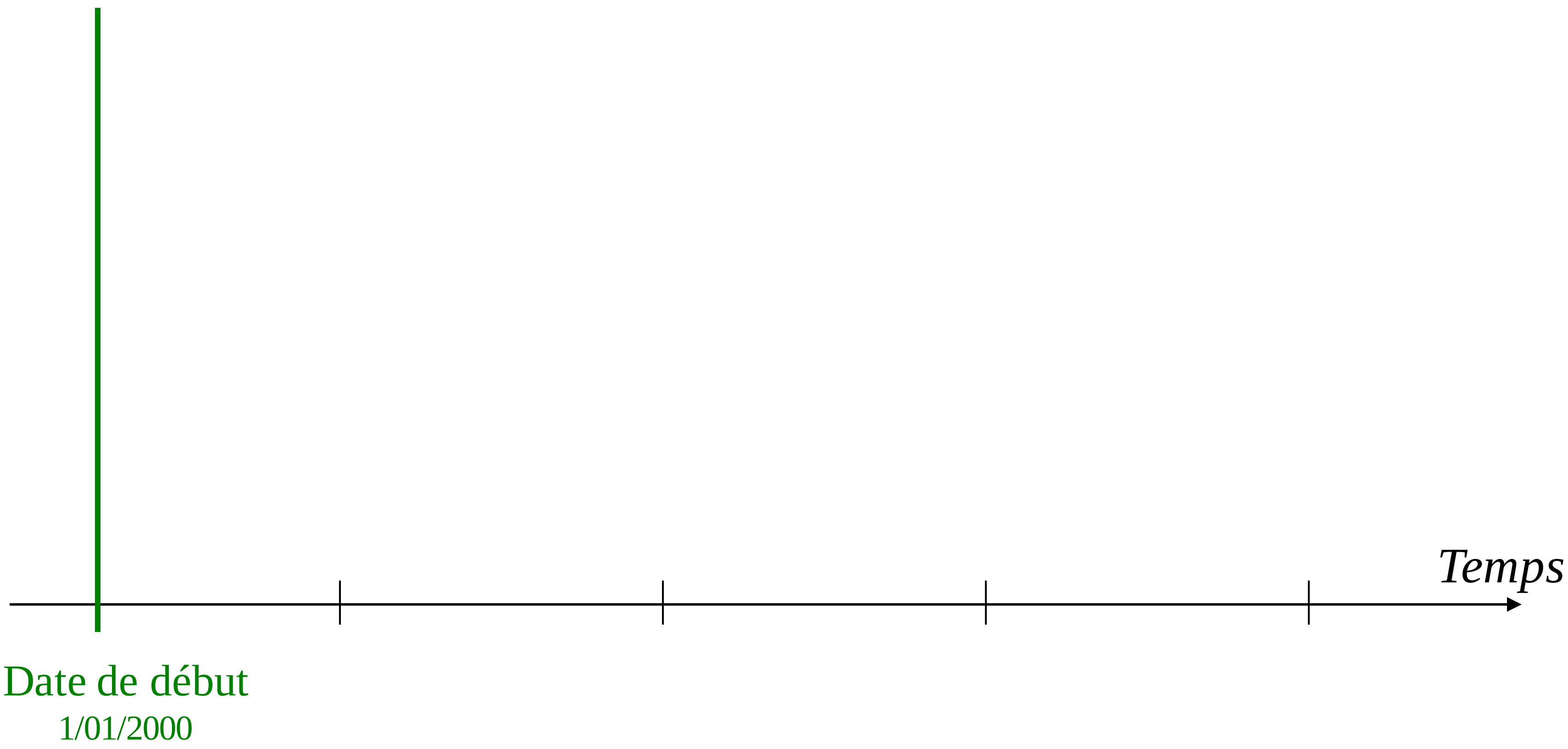
$$\frac{VV}{VV + DC}$$

Définitions (1)

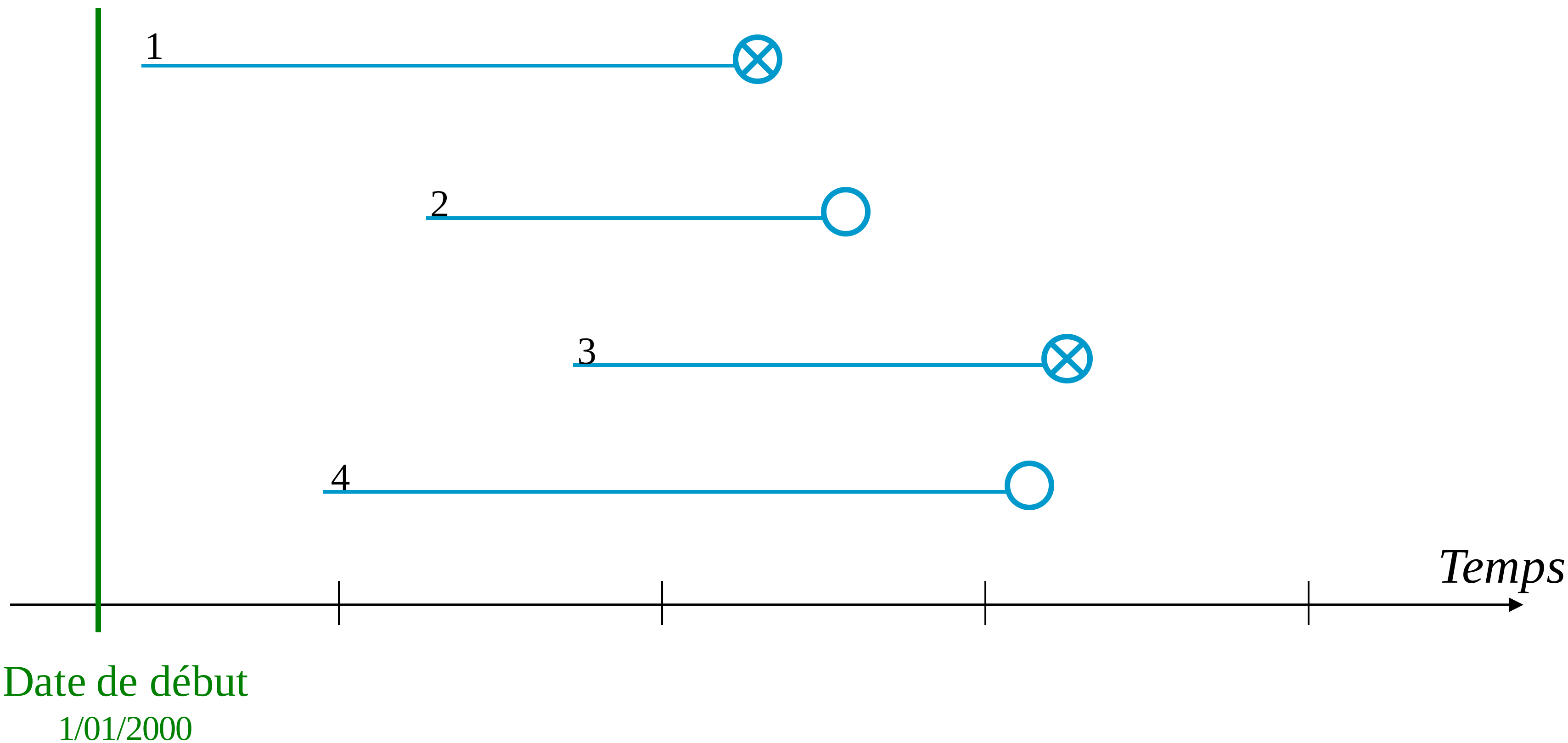
- Fonction de survie : $S(t)=Pr(T \geq t)$
- $S(0)=1$, $S(t)$ est décroissante et tend vers 0 quand $t \rightarrow \infty$



Définitions (2)

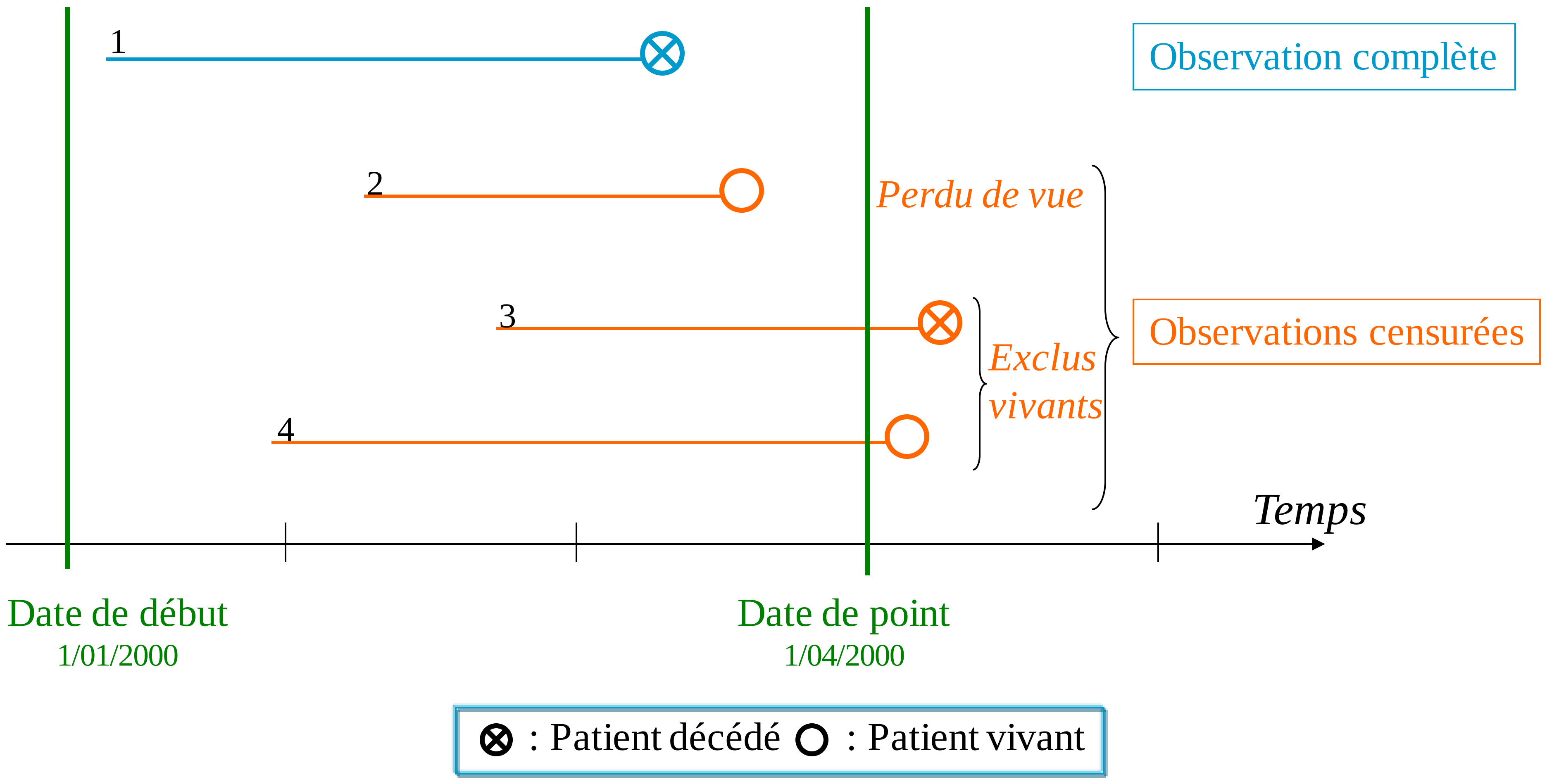


Définitions (3)

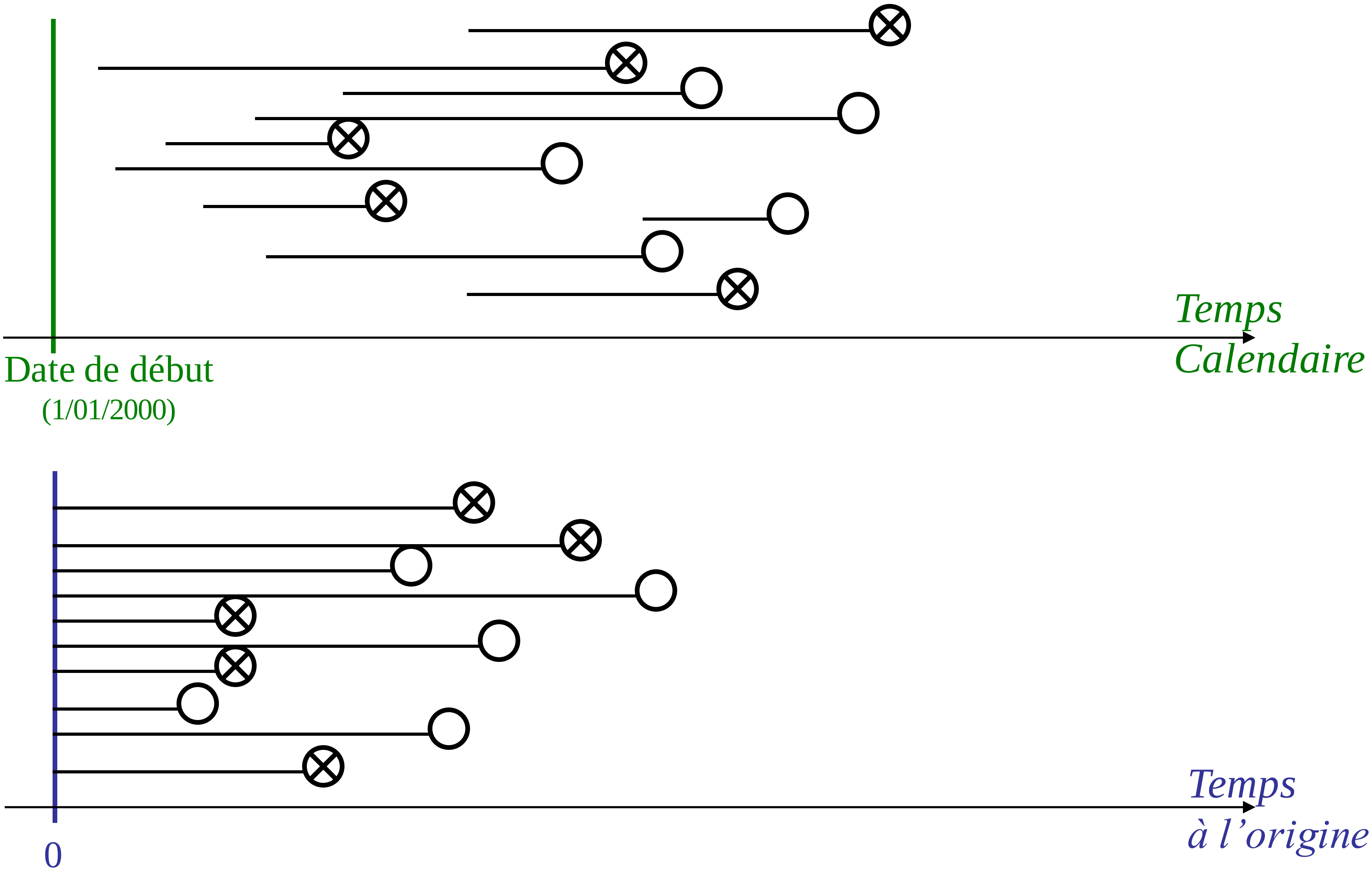


⊗ : Patient décédé ○ : Patient vivant

Définitions (4)



Notion d'Origine



Censures (1)

- Une observation est censurée (à droite) si on sait que le temps de suivi du sujet *i* est inférieur à la variable aléatoire du temps de survie de ce sujet *i*

Censure à droite : $T_i > t_i$

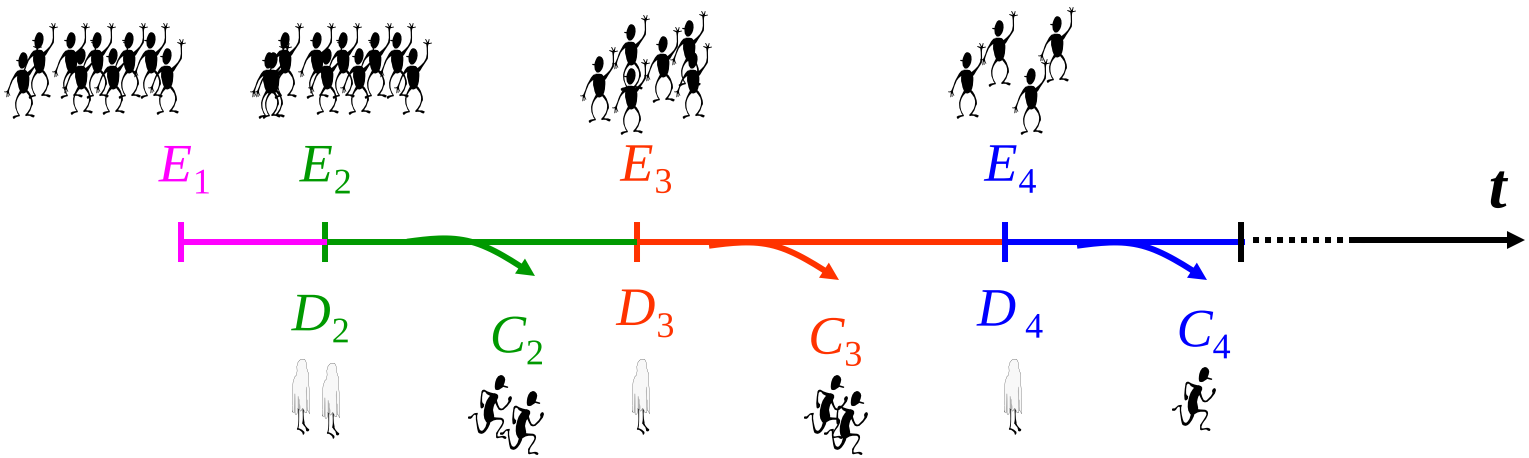
État aux dernières nouvelles : $\delta_i = 0$

Censures (2)

- Non aléatoire
  - ✓ Informative
    - Exemple : expérience de durée prédéterminée, survenue d'un autre événement empêchant l'observation de l'évènement étudié
  - ✓ Entraine des biais (estimation, comparaison) si non prise en compte
- Aléatoire
  - ✓ Non informative
  - ✓ Hypothèse de base des modèles de survie
    - Indépendance entre la cause de la censure et l'évènement étudié (*aucune variable n'affecte simultanément la cause de censure et l'évènement étudié*)

Méthode de Kaplan-Meier (1)

- Nombre de patients exposés au risque de décès (*E*) :

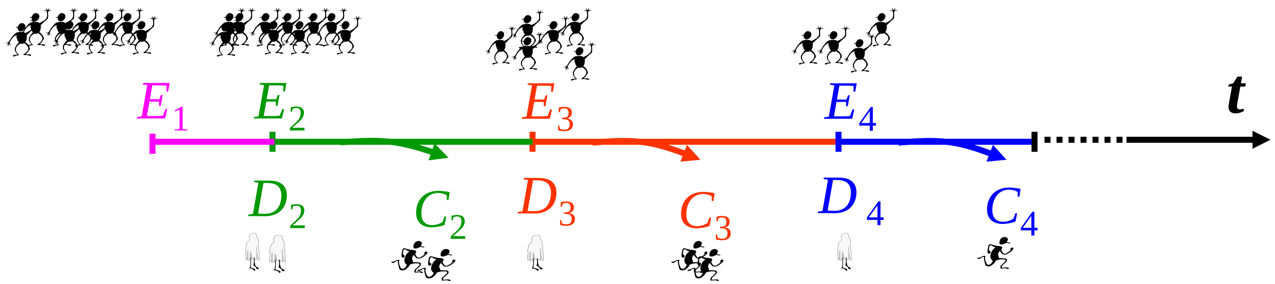


$$\boxed{E_3 = E_2 - D_2 - C_2}$$

C_2 : nombre de censures (exclus vivants, perdus de vue) pendant le deuxième intervalle de temps

D_2 : nombre de décès au début du deuxième intervalle de temps

Méthode de Kaplan-Meier (2)



- Taux de survie conditionnelle :

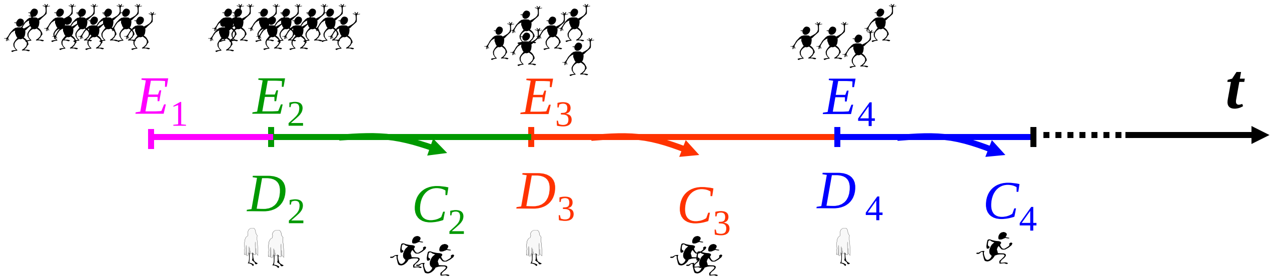
$$\boxed{S(t_4 \mid t_3) = \frac{E_3 - D_3}{E_3}}$$

avec $E_3 = E_2 - D_2 - C_2$

D'une manière générale on a :

$$\boxed{S(t_{i+1} \mid t_i) = \frac{E_i - D_i}{E_i}}$$

Méthode de Kaplan-Meier (3)



- Taux de survie à un moment donné :

$$S(t_3) = \frac{E_1}{E_1} \times \frac{E_2 - D_2}{E_2} \times \frac{E_3 - D_3}{E_3}$$

D'une manière générale on a :

$$S(t) = \frac{E_2 - D_2}{E_2} \times \frac{E_3 - D_3}{E_3} \times \ldots \times \frac{E_t - D_t}{E_t}$$

Estimation de la Variance

$$S(t) = \frac{E_2 - D_2}{E_2} \times \frac{E_3 - D_3}{E_3} \times \ldots \times \frac{E_t - D_t}{E_t} = \sum_{t_i \le t} \frac{(E_i - D_i)}{E_i}$$

- Formule de Greenwood

$$\boxed{\mathrm{Var}(S(t)) = [S(t)]^2 \sum_{t_i \le t} \frac{D_i}{E_i (E_i - D_i)}}$$

Méthode de Kaplan-Meier : Exemple (1)

- Exemple : durées de rémission (semaine) chez des patients ayant une leucémie aiguë et traités par 6-MP (Freireich, 1963)

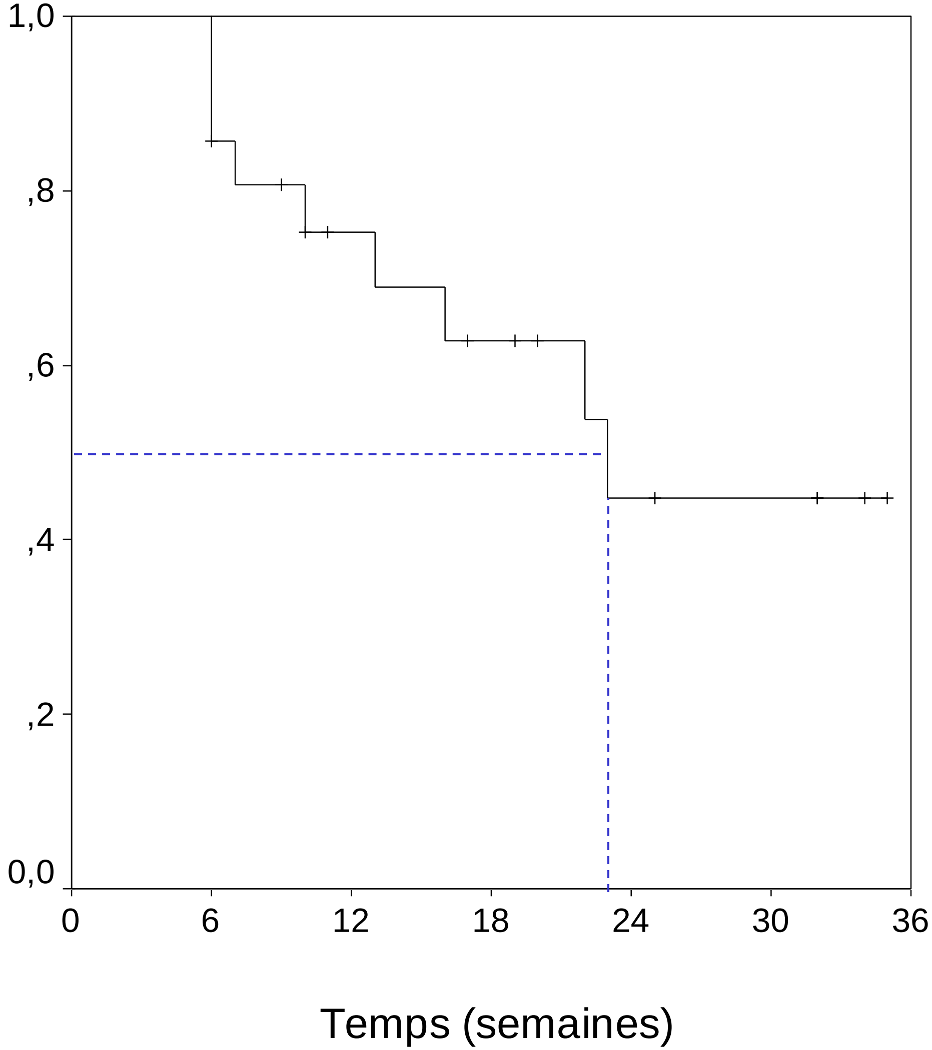
  Data : 6, 6, 6, 6*, 7, 9*, 10, 10*, 11*, 13, 16, 17*, 19*, 20*, 22, 23, 25*, 32*, 32*, 34*, 35*

  * : données censurées (pas de rechute aux dernières nouvelles)

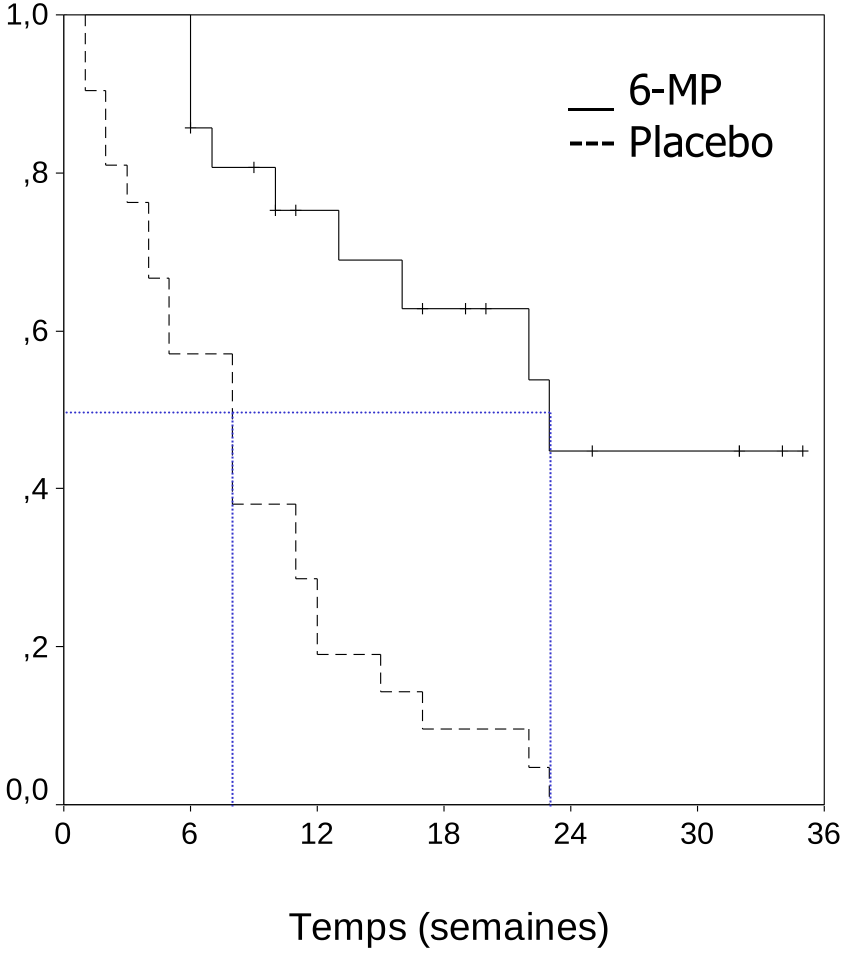
Méthode de Kaplan-Meier : Exemple (2)

Temps	Rechutes	Censures $[t_{i-1}, t_i[$	Exposés t_i	Survie Conditionnelle en t_i	Survie juste après t_i
6	3	0	21	18/21	0,857
7	1	1	17	16/17	0,807
10	1	1	15	14/15	0,753
13	1	2	12	11/12	0,690
16	1	0	11	10/11	0,627
22	1	3	7	6/7	0,538
23	1	0	6	5/6	0,448

Méthode de Kaplan-Meier : Exemple (3)



Comparaison de la Survie dans 2 Groupes



- Survie moyenne
- Survie médiane
- Survie à t
- Distribution des durées de survie

Test du logrank (1)

- Basé sur les rangs des observations
- H_0 : les taux de décès attendus dans les 2 groupes sont identiques
- Principe :
  - ✓ Calculer pour chaque date de décès le nombre attendu de décès qui serait observé si les taux de décès étaient les mêmes dans les 2 groupes
  - ✓ Cumuler les différences entre observés et attendus jusqu'à la date du dernier décès observé
  - ✓ Évaluer la signification de la différence cumulée en la comparant à son erreur type

Test du logrank (2)

- Ordonner les dates de décès
- Pour chaque date de décès

	Groupe 1	Groupe 2	Total
Décès	d_{1i}	d_{2i}	d_{+i}
Effectif	n_{1i}	n_{2i}	n_{+i}

- Sous H_0 : $e_{2i} = d_{+i}.n_{2i}/n_{+i}$
- La variance de $\Delta_i = d_{2i} - e_{2i}$ est :

$$Var(\Delta_i) = \frac{n_{1i} n_{2i} d_{+i} (n_{+i} - d_{+i})}{n_{+i}^2 (n_{+i} - 1)}$$

- Le test est : $\left(\sum_i \Delta_i\right)^2 \Big/ \sum_i Var(\Delta_i) \sim \chi^2$

Exemple (1)

Groupe	Temps de suivi (* censure)
Placebo	1, 1, 2, 2, 3, 4, 5, 5, 8, 8, 8, 8, 11, 11, 12, 12, 15, 17, 22, 23
6-MP	6, 6, 6, 6*, 7, 9*, 10, 10*, 11*, 13, 16, 17*, 19*, 20*, 22, 23, 25*, 32*, 32*, 34*, 35*

(Freireich, 1963)

Exemple (2)

Temps	n1	d1	n2	d2	n+	d+	e2	var
1	21	0	21	2	42	2	1,00	0,49
2	21	0	19	2	40	2	0,95	0,49
3	21	0	17	1	38	1	0,45	0,25
4	21	0	16	2	37	2	0,86	0,48
5	21	0	14	2	35	2	0,80	0,47
6	21	3	12	0	33	3	1,09	0,65
7	17	1	12	0	29	1	0,41	0,24
8	16	0	12	4	28	4	1,71	0,87
10	15	1	8	0	23	1	0,35	0,23
11	13	0	8	2	21	2	0,76	0,45
12	12	0	6	2	18	2	0,67	0,42
13	12	1	4	0	16	1	0,25	0,19
15	11	0	4	1	15	1	0,27	0,20
16	11	1	3	0	14	1	0,21	0,17
17	10	0	3	1	13	1	0,23	0,18
22	7	1	2	1	9	2	0,44	0,30
23	6	1	1	1	7	2	0,29	0,20
Total		9		21			10,75	6,26

Test du logrank = $(21 - 10{,}75)^2/6{,}26 = 16{,}79$

Test du logrank (3)

- Extensions du test
  - ✓ Comparaisons des distributions de survie de k groupes
  - ✓ Stratification sur des facteurs de confusion