

## I. Introduction

### ① - Problématique

On parle dans ce cours de la durée écoulée jusqu'à la survenance d'un événement donné. L'exemple le plus commun est l'étude de la durée de vie, ce qui amène à définir l'événement d'intérêt comme la mort (d'un composant, d'un individu, ...). Une des difficultés réside dans la bonne compréhension et définition de l'événement d'intérêt, ce qui peut parfois conduire à des erreurs d'interprétation et donc de mise en pratique de cette théorie.

### ② - Quelques illustrations/exemples:

Les modèles de durée (duration models), ou l'analyse de survie (survival analysis) ou time-to-event data), sont une branche des statistiques largement utilisée dans de nombreux domaines. On peut citer par exemple :

- le domaine biomédical: des durées de survie sont étudiées dans le contexte d'études longitudinales comme les enquêtes de cohorte (suivi de patients dans le temps) par exemple. On étudie par exemple si l'événement correspondant à l'apparition d'une maladie (par exemple, le temps avant une rechute), une durée de guérison;

- la fiabilité : on s'intéresse à la durée de vie de composants, par exemple électroniques, ou la durée de fonctionnement d'une machine plus généralement
- la démographie : pour l'étude de l'espérance de vie, des projections sur la pyramide des âges, Ici, l'événement d'intérêt est le décès des individus;
- l'actuarial : on s'intéresse par exemple au temps écoulé entre deux sinistres, à la construction de table de mortalité / longévité, à la durée en incapacité / invalidité.

→ Dans tous ces exemples, on réalise que la particularité des données observées (celles qui vont nous servir à faire des estimations statistiques) réside dans le fait que l'échantillon étudié comporte des observations pour lesquelles l'événement d'intérêt a été observé, et des observations pour lesquelles ce n'est pas encore le cas!

### ③ - Vocabulaire

Introduisons un certain nombre de termes couramment utilisés lors de la manipulation de modèles de durée.

- Date d'origine : c'est l'origine de la durée étudiée. Selon le contexte, on peut penser à :
  - une date de naissance;
  - une date d'opération (chirurgicale), ou de début de maladie;
  - une date d'entrée dans une étude;
  - une date de début d'exposition à un facteur de risque...
- Date de point : date au-delà de laquelle on arrête l'étude et on ne collecte plus d'information sur les individus/sujets.



- Date des dernières nouvelles: date la plus récente pour laquelle on dispose d'informations sur les sujets. ②

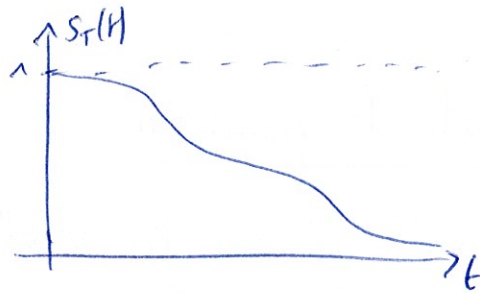
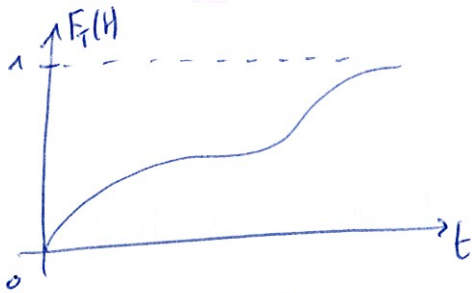
## II - Notions probabilistes fondamentales.

Dans toute la suite du cours, on note  $T$  la durée de survie. On considère que  $T$  est une variable aléatoire positive ou nulle, absolument continue.

On peut caractériser la loi de  $T$  de plusieurs manières équivalentes, citées ci-après. (Toutes ces notions se déduisent l'une de l'autre, et sont en ce sens équivalentes).

### ① - La fonction de survie (survival function)

La fonction de survie, est, comme son nom l'indique, la probabilité de survie jusqu'à l'instant  $t$  (où  $t$  est fixé). C'est donc  $S(t) = P(T > t)$ ,  $t \geq 0$ .



C'est le complémentaire à 1 de la fonction de répartition, au sens où :

$$S(t) = P(T > t) = 1 - P(T \leq t) = 1 - F(t).$$

### ② - La fonction de répartition (cumulative distribution function)

Très bien connu en probabilité, elle représente donc la probabilité de "mourir" avant l'instant  $t$ . Par définition, elle vaut  $F(t) = P(T \leq t)$ .

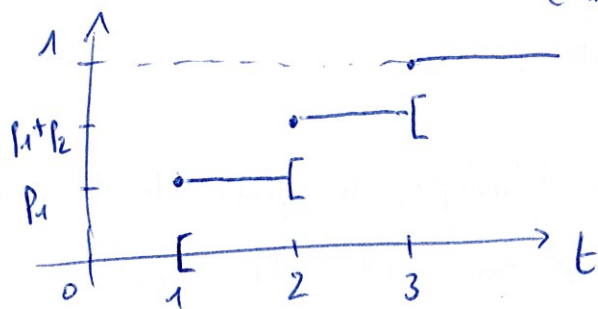
Remarque: si la loi de  $T$  est continue, on sait que  $P(T=t)=0$ .

Ainsi, on peut définir  $S(t) = P(T > t)$  ou  $S(t) = P(T \geq t)$ , et cela n'a aucune incidence. Lorsque  $F$  comporte des sauts (si le temps est discret, compté en semaine/mois/année par exemple), on introduit les notations:

$$F^-(t) = P(T < t) \quad \text{et} \quad F^+(t) = P(T \leq t).$$

Ainsi  $F^-$  est la limite à gauche et  $F^+$  la limite à droite de  $F$ .

On rappelle qu'une fonction de répartition est une fonction càdàg: par exemple, si  $T$  est à valeurs dans  $\{1, 2, 3\}$  avec  $\begin{cases} P(T=1) = p_1 \\ P(T=2) = p_2 \\ P(T=3) = 1 - p_1 - p_2 \end{cases}$ , alors la FDR est illustrée comme suit:



### ③ - La densité de probabilité

En admettant que  $T$  est une loi admettant une densité, la densité de probabilité  $f(t) \geq 0$  est la fonction telle que pour tout  $t \geq 0$ :

$$F(t) = \int_0^t f(u) du.$$

Ainsi,  $f(t)$  est la dérivée de la fonction de répartition au point  $t$ . On peut

$$\begin{aligned} \text{ainsi écrire} \quad f(t) = F'(t) &= \lim_{h \rightarrow 0} \frac{F(t+h) - F(t)}{t+h-t} = \lim_{h \rightarrow 0} \frac{1}{h} (P(T \leq t+h) - P(T \leq t)) \\ &= \lim_{h \rightarrow 0} \frac{P(t < T \leq t+h)}{h} \end{aligned}$$

$\Rightarrow$  Proba de "mourir" dans un petit intervalle de temps après l'instant  $t$ .



Étant donné que  $F(t) = 1 - S(t)$ , on a  $F'(t) = -S'(t) = -\frac{d}{dt} S(t)$ . ③

Ainsi, on a aussi :  $f(t) = -S'(t)$ .

#### ④ - Le taux d'incidence, ou risque instantané

Pour  $t$  fixé, le taux de hasard (hazard rate) ou risque instantané représente la probabilité de "mourir" dans un petit intervalle de temps après  $t$ , sachant qu'on a "survécu" jusqu'en  $t$ . Ainsi, on peut l'écrire :

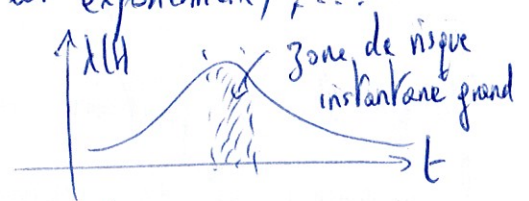
$$\lambda(t) = \lim_{h \rightarrow 0} \frac{P(t < T \leq t+h | T > t)}{h} \stackrel{\text{Bayes}}{=} \lim_{h \rightarrow 0} \frac{P(t < T \leq t+h, T > t)}{h \times P(T > t)}$$

$$= \frac{1}{P(T > t)} \lim_{h \rightarrow 0} \frac{P(t < T \leq t+h)}{h} = \frac{1}{S(t)} f(t) = \frac{f(t)}{S(t)} = -\frac{S'(t)}{S(t)}$$

On peut déduire de cette dernière expression que  $\lambda(t) = -\frac{d}{dt} (\ln S(t))$

⇒ le taux de hasard peut être croissant, décroissant, constant... Il est très utile pour caractériser des phénomènes tels que l'usure (taux croissant), l'absence de mémoire d'un phénomène (constant, exemple de la loi exponentielle), ...

#### ⑤ - Le taux de hasard cumulé



A travers la définition du taux de hasard  $\lambda$ , on constate que  $\lambda(t) \geq 0, \forall t$ . Comme son nom l'indique, le taux de hasard cumulé est l'intégrale de  $\lambda$  :

$$\Lambda(t) = \int_0^t \lambda(u) du = \int_0^t -\frac{1}{S(u)} \ln S(u) du = -\ln(S(t)).$$

On déduit notamment:

- $S(t) = e^{-\Lambda(t)} = e^{-\int_0^t \lambda(u) du}$
- $f(t) = \lambda(t) S(t) = \lambda(t) e^{-\int_0^t \lambda(u) du}$

Conclusion: on voit bien que disposer d'une de ces fonctions permet de retrouver l'expression des autres fonctions.

### III - Quantités utiles

#### ① - Moyenne et variance de la durée de vie

On rappelle les définitions usuelles de ces quantités:

- $E[T] = \int_0^{+\infty} t f(t) dt \stackrel{\text{IPP}}{=} \int_0^{+\infty} (1-F(t)) dt = \int_0^{+\infty} S(t) dt$

- $\text{Var}(T) = E[(T-E[T])^2] = E[T^2] - (E[T])^2$   
 $\stackrel{\text{IPP}}{=} 2 \int_0^{+\infty} t S(t) dt - (E[T])^2$

Vo que les 5 fonctions vues précédemment sont équivalentes, on peut déduire l'espérance et la variance à partir de l'une d'entre elles.

⚠: la réciproque est fausse! Notamment, deux variables aléatoires différentes peuvent avoir la même moyenne, variance, ...!



## ② - Quantiles

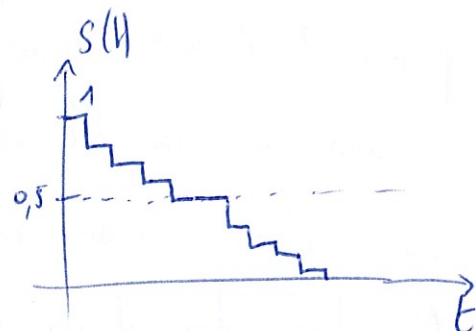
(4)

On s'intéresse souvent au quantile à 50%, soit la médiane.

La médiane de la durée de survie est le temps  $t$  pour lequel la probabilité de survie  $S(t)$  est égale à 0,5.

C'est donc la valeur  $t_m$  telle que  $S(t_m) = 0,5$ .

- Si l'on dispose d'une fonction de survie en escalier (telle que celle donnée par l'estimateur de Kaplan-Meier),



l'interprétation est compliquée... (beaucoup de valeurs de  $t$  correspondent à la médiane!).

- On peut donner un intervalle de confiance de ce temps médian: pour cela, on a besoin de disposer d'un intervalle de confiance de la fonction de survie (noté  $[B_i, B_s]$ ), puis on applique la fonction inverse: cela donne  $[S^{-1}(B_s); S^{-1}(B_i)]$ .

⚠: permutation des bornes puisque  $S$  est décroissante...

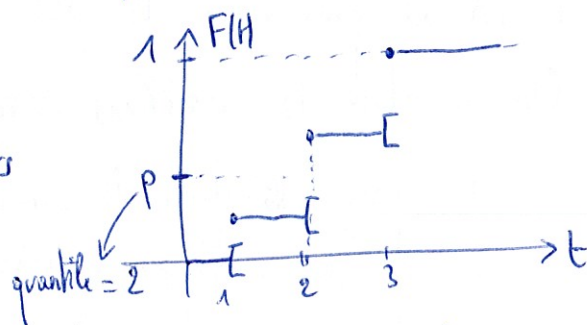
- De manière générale, la fonction quantile de la durée de survie est définie par:

$$q(p) = \inf \{ t : F(t) = P(T \leq t) \geq p \}, \quad 0 < p < 1.$$

$$= \inf \{ t : S(t) \leq 1 - p \}.$$

Si la FDR  $F$  est strictement  $\uparrow$  et continue, alors

$$q(p) = F^{-1}(p) = S^{-1}(1-p).$$



$\Rightarrow$  Le quantile  $q(p)$  représente le temps où une proportion  $p$  de la population a disparu.  
(au moins!)

## IV - Censure, Troncature

Comme énoncé en introduction, la plus grande particularité des données liées à l'observation d'un phénomène de durée avant un événement d'intérêt est la présence de données "incomplètes". En effet, pour certains sujets/individus, l'événement n'aura pas encore été observé... C'est ce qui est couramment appelé la censure.

Pour d'autres, on n'aura même rien observé du tout: c'est la troncature.

On dispose donc d'une information partielle, au sens où l'on n'observe pas en pratique que des réalisations de la variable aléatoire  $T$ . Comme l'on n'observe pas des répliques iid. de durée  $T$  (mais plutôt une version perturbée de  $T$ , perturbée par les processus de censure et de troncature), on ne peut pas appliquer tel quel les résultats de la statistique classique (comme la loi des grands nombres par exemple). Donnons maintenant davantage de détail sur les processus de censure et troncature.

### ① - La censure

C'est le phénomène le plus courant quand on travaille avec des données de survie.

On introduit les notations suivantes: pour l'individu  $i$ ,

- $T_i$  désigne la durée de survie
- $C_i$  désigne le temps de censure
- $X_i$  désigne la durée observée.

Il existe la censure à droite (la plus commune), la censure à gauche, ainsi que la censure par intervalle.

Il y a même plusieurs types de censure pour la censure droite par exemple.



## a) Censure à droite:

Si l'événement d'intérêt n'a pas été observé à la dernière observation de l'individu, on dit que l'individu est censuré à droite. Ainsi, les durées observées ne le sont pas toutes totalement: pour certaines, on sait seulement qu'elles seraient supérieures à une certaine valeur (connue).

### \* Censure de type I:

C'est la plus classique. Soit  $C$  une valeur fixée (de durée).

Au lieu d'observer  $T_1, T_2, \dots, T_n$ , on n'observe  $T_i$  que si  $T_i \leq C$ .

Si on sait que  $T_i > C$ . Ainsi, la durée observée  $X_i$  est donnée par:

$$X_i = T_i \wedge C = \min(T_i, C).$$

Ce type de censure se rencontre souvent dans les applications industrielles, par exemple lorsque l'on teste la fiabilité des objets (on les observe sur une certaine durée).

### \* Censure de type II:

Contexte où l'on observe les durées de vie de  $n$  individus jusqu'à ce que l'on observe l'événement pour  $k$  d'entre eux. On arrête alors l'étude.

Introduisons  $T_{(i)}$  et  $X_{(i)}$  les durées de vie et les durées observées ordonnées.

La date de censure vaut donc  $T_{(k)}$ , et on observe:

$$\left\{ \begin{array}{l} X_{(1)} = T_{(1)} \\ X_{(2)} = T_{(2)} \\ \vdots \\ X_{(k)} = T_{(k)} \\ X_{(k+1)} = T_{(k)} \\ \vdots \\ X_{(n)} = T_{(k)} \end{array} \right\} \text{ ne change plus!}$$

## \* Censure de type III:

On l'appelle aussi censure aléatoire de type I.

Introduisons  $C_1, C_2, \dots, C_n$  des variables aléatoires i.i.d. - Ici, on observe:

$$X_i = T_i \wedge C_i = \min(T_i, C_i).$$

On résume l'information disponible par:

- la durée réellement observée  $X_i$
- l'indicatrice  $\delta_i$ : elle vaut 1 lorsque l'événement d'intérêt a été observé.

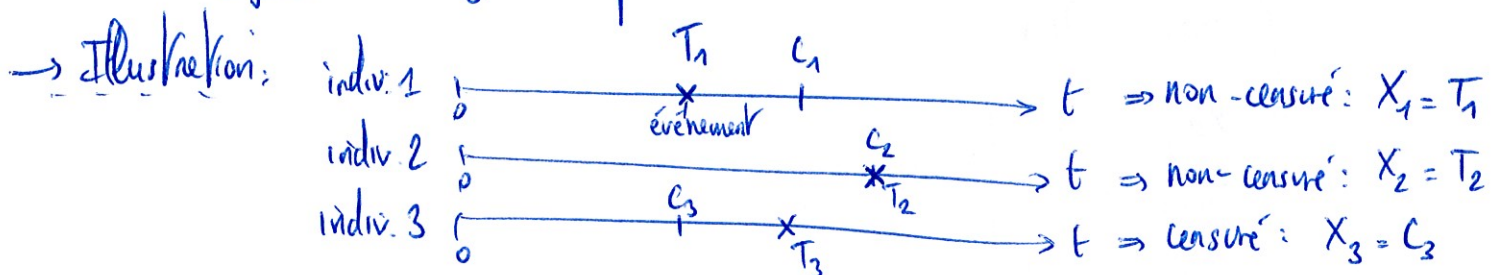
$$\text{Ainsi: } \delta_i = \mathbb{1}_{\{T_i \leq C_i\}} -$$

Si  $\delta_i = 0$ , on dispose donc d'une donnée incomplète, en l'occurrence une durée censurée à droite.

Cette censure est la plus courante, car elle correspond le mieux à la réalité du terrain - En particulier, elle permet de rendre compte de: (lors d'un essai thérapeutique)

- un arrêt / changement de traitement: exclusion de patients à cause d'effets secondaires ou d'inefficacité du traitement. (imprévu au départ).
- les "perdus de vue": ce sont les individus sortant de l'étude à une durée imprévisible aléatoire (déménagement, ...).
- la fin de l'étude: l'étude se termine et des patients sont encore "vivants".

Ces "exclus-vivants" sont censurés comme les "perdus-de-vue", mais les mécanismes sous-jacents sont très différents... En particulier, la censure peut être informative chez les "perdus-de-vue".





→ Remarque fondamentale : dans toute la suite du cours, on supposera que  $T_i$  est indépendant de  $C_i$ . Autrement dit, la censure et l'événement sont  $\perp$ . ⑥  
Cette hypothèse est très pratique d'un point de vue calculatoire, et simplifie grandement la modélisation statistique. En revanche, il faut tout de même vérifier qu'elle ne soit pas totalement fautive en fonction du contexte...

• Exemple 1 : cette hypothèse est naturelle quand on parle d'un démençement ayant causé la censure.

• Exemple 2 : elle est nettement plus discutable si l'on suit un patient d'une étude à cause d'un état de santé critique ; en effet la censure apporte une information sur l'état de santé du patient, et donc sur l'événement d'intérêt...

### b) Censure à gauche :

C'est le cas où l'individu a déjà subi l'événement avant d'être observé.

⇒ On sait donc que la date de l'événement est inférieure à une certaine date (connue). Ainsi, pour chaque individu, on associe un couple de v.a.  $(X, \delta)$  t.p. :

$$\begin{cases} X = \max(T, C) = TVC \\ \delta = \mathbb{1}_{\{T \geq C\}} : \delta = 1 \text{ si on observe l'événement d'intérêt.} \end{cases}$$

On considère aussi ici  $T \perp C$ .

• Exemple : on s'intéresse à l'heure à laquelle des animaux vont s'alimenter. Le temps d'événement est observé si l'animal y va après l'arrivée des observateurs. On observe ainsi le maximum entre l'heure d'arrivée des observateurs et l'heure de repos des animaux.

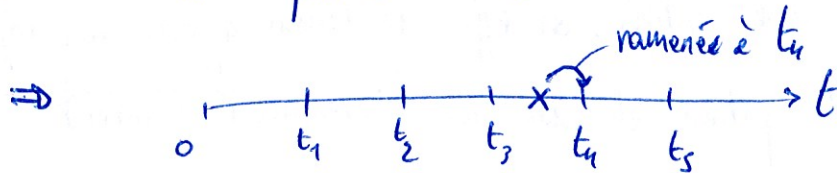
→ La censure à gauche est très peu fréquente...

→ Donner des exemples / proposer.

### c) Censure par intervalle:

On parle de censure par intervalle lorsque l'on n'observe pas exactement le temps de l'événement, mais l'on sait qu'il a lieu (ou non!) entre deux dates. Cette censure apparaît souvent lorsque l'on suit une cohorte, car les mesures sont prises par intermittence --- On sait alors uniquement que l'événement s'est produit entre deux temps d'observation.

• Rq: pour simplifier en pratique, on prend souvent le temps d'événement égal au temps de la dernière visite pour se ramener à de la censure à droite!



### ② - La Troncature.

La notion de troncature est radicalement différente de celle de censure. En effet, la troncature ne renvoie à l'échantillonnage lui-même.

Une v.a.  $T$  est tronquée par un sous-ensemble (éventuellement aléatoire)  $A \in \mathbb{R}^+$  si au lieu d'observer  $T$ , on n'observe  $T$  que si  $T \in A$ .

⇒ Les points de l'échantillon tronqué appartiennent tous à  $A$ . ⇒ ils suivent donc la loi de  $T$  conditionnée par l'appartenance à  $A$ .

⇒ Une partie des points  $T$  n'est donc pas du tout observée, sans aucune information.

⇒ On n'étudie qu'un sous-échantillon!



### a) La Troncature gauche:

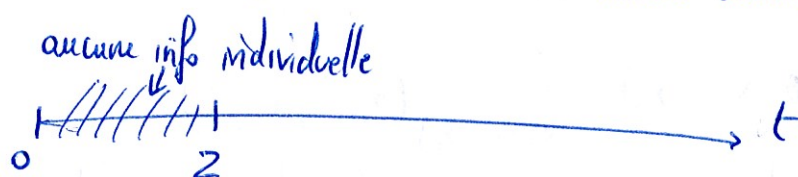
(7)

La plus courante. Introduisons une v.a.  $Z$ , indépendante de  $T$ .

On parle de troncature gauche quand  $T$  n'est observable que si  $T > Z$ .

On observe le couple  $(T, Z)$  avec  $T > Z$ .

→ Exemple: en étudiant la durée de vie d'une population depuis une cohorte tirée au hasard dans cette population, on n'étudiera que la survie des individus vivants à l'inclusion de l'étude... (on omet tous ceux morts avant).



b) La Troncature droite: c'est le même raisonnement, sauf que  $T$  n'est observable que si  $T < Z$ .

→ Exemple:

c) Troncature par intervalle: on parle de troncature par intervalle lorsque une durée est tronquée à gauche et à droite.

→ Exemple: étude des patients d'un registre: les patients diagnostiqués avant la mise en place du registre ou répertoriés après la consultation du registre ne seront pas inclus dans l'étude.

## V - la vraisemblance

Rappelons que nous considérons que  $T \perp C$ .

On note  $f$  la densité de probabilité de  $T$ , et  $g$  celle de  $C$ .

La fonction de survie de  $T$  est notée  $S$ , celle de  $C \vee G$ . Admettons également que la distribution de  $T$  est une loi dépendante d'un paramètre de dimension finie.

### ① - La vraisemblance dans le cas de données complètes.

En l'absence de censure/troncature, la vraisemblance se définit simplement comme la densité multivariée associée à un échantillon. Ainsi, si l'on observe l'échantillon  $x = (x_1, x_2, \dots, x_n)$  réalisations de  $X_1, X_2, \dots, X_n$  de lois iid, avec densité commune  $f_{X_i}$  de paramètre  $\theta$ , on écrit généralement:

$$L(\theta; x) = L(\theta; x_1, \dots, x_n) = \prod_{i=1}^n f_{X_i}(x_i; \theta).$$

En se servant de l'indépendance entre les  $X_i$ , il vient par simplification avec un échantillon  $x$ :

$$L(\theta; x) = \prod_{i=1}^n f_{X_i}(x_i; \theta).$$

Comme le but est généralement de maximiser cette fonction par rapport à la valeur  $\hat{\theta}$  de  $\theta$  la plus probable sachant l'échantillon observé  $x$ , on a tendance à considérer la log-vraisemblance  $l(\theta; x) : l(\theta; x) = \ln L(\theta; x) = \sum_{i=1}^n \ln(f_{X_i}(x_i; \theta))$ .

$\Rightarrow$  On constate que la vraisemblance de l'échantillon est constituée de  $n$  termes de vraisemblances individuelles, chacun apportant une contribution à la vraisemblance globale.

### ② - Cas des données de survie (censurées à droite).

On rappelle que toute l'information est ici contenue dans le couple  $(X_i, \delta_i)$ , où

$X_i = \min(T_i, C_i)$  est la durée observée et  $\delta_i = \mathbb{1}_{\{T_i \leq C_i\}}$  est l'indicateur de censure (qui vaut 1 lorsque l'événement d'intérêt a été observé). Ainsi, la  $i^{\text{e}}$  contribution à la vraisemblance,  $L_i$ , peut être vue comme celle d'une loi de Bernoulli associée à la censure.



$$\begin{aligned} L_i &= \underset{\text{v.e. observé}}{P(X_i \in [t_i, t_i+dt])}^{1-\delta_i+\delta_i}_{\text{v.e. observé}} \\ &= P(X_i \in [t_i, t_i+dt])^{\delta_i} \times P(X_i \in [t_i, t_i+dt])^{1-\delta_i} \\ &= P(T_i \in [t_i, t_i+dt])^{\delta_i} \times P(X_i \in [t_i, t_i+dt])^{1-\delta_i} \\ &= P(T_i \in [t_i, t_i+dt], T_i \leq C_i; \theta)^{\delta_i} \times P(C_i \in [t_i, t_i+dt], T_i > C_i; \theta)^{1-\delta_i} \\ &\stackrel{T \perp C}{=} \left[ f(t_i; \theta) G(t_i) \right]^{\delta_i} \times \left[ g(t_i) S(t_i; \theta) \right]^{1-\delta_i} \end{aligned}$$

Comme la censure est considérée non-informative ici, le paramètre  $\theta$  n'apparaît pas dans la loi de  $C$ . Étant donné que l'on optimise cette vraisemblance en  $\theta$ , cela veut dire qu'on peut considérer une version simplifiée de cette contribution...

Ainsi, on prendrait  $L'_i = \underset{\theta_i}{f(t_i; \theta)}^{\delta_i} S(t_i; \theta)^{1-\delta_i}$

La vraisemblance pourrait alors s'écrire : 
$$L = \prod_{i=1}^n L'_i = \prod_{i=1}^n f(t_i; \theta)^{\delta_i} S(t_i; \theta)^{1-\delta_i}$$

- La présence de données censurées à droite est mise en évidence dans  $L$  à travers le terme faisant apparaître la fonction de survie: en effet, la seule info. disponible pour les observations censurées est que la durée de vie est plus grande!
- En considérant uniquement les données non-censurées, on aurait  $L = \prod_{i=1}^n f(x_{ii}; \theta)^{\delta_i}$ . Cela conduirait à un estimateur  $\hat{\theta}$  biaisé asymptotiquement. En particulier, on aurait tendance à la fin à sous-estimer les durées de vie.
- L'expression de  $L$  ne change pas par une censure droite non-aléatoire.
- Si les observations  $T_i$  ne sont pas iid (cas de modèles de régression par exemple), on peut toujours utiliser cette expression en indiquant par  $i$  la densité  $f$  et la survie  $S$ .

### ③ - Cas de la censure à droite de type II.

On rappelle que cette censure concerne le cas où le temps de censure est défini par le  $k^{\text{ième}}$  événement d'intérêt dans une population de  $n$  individus. Naturellement, la vraisemblance s'écrit alors :

$$L = \prod_{i=1}^k C_n^k f(x_i; \theta) S(x_k; \theta)^{n-k} = \frac{n!}{k!(n-k)!} \prod_{i=1}^k f(x_i; \theta) S(x_k; \theta)^{n-k}$$

après avoir ordonné les  $X_i$  par ordre croissant.

On reconnaît la structure d'une vraisemblance de loi binomiale, où le nombre de combinaisons de  $k$  événements parmi  $n$  individus intervient directement.

### ④ - Données longitudinales électoralement à gauche.

les v.a.  $T_i$  sont soumises à truncature par les v.a.  $Z_i$ , avec  $T_i \perp Z_i$ .

On dispose d'un échantillon  $(T_i, Z_i)_{i=1, \dots, N}$  avec  $N$  aléatoire : en effet, on observe uniquement les individus qui vérifient  $T_i \geq Z_i$  dans une population inconnue de taille  $n$ . On obtient ainsi :

$$L = \prod_{i=1}^N P(T_i \in [x_i, x_i + dx] | T_i \geq z_i) = \prod_{i=1}^N \frac{P(T_i \in [x_i, x_i + dx], T_i \geq z_i)}{P(T_i \geq z_i)} = \prod_{i=1}^N \frac{f(x_i)}{S(z_i)}$$

→ Cette vraisemblance est conditionnelle par rapport à  $N$ , et aux valeurs  $z_i$ .

→ On peut aussi écrire la vraisemblance comme suit, puisque  $x_i \geq z_i$  par construction :

$$L' = \prod_{i=1}^N L(T_i, z_i | T_i \geq z_i) = \frac{1}{P(T \geq z)^N} \prod_{i=1}^N L(T_i, z_i)$$

expression utile si on a une opinion sur la loi de  $Z$

$$= \left[ \int \int \mathbb{1}_{t \geq z} f(z) f(t) dz dt \right]^{-N} \prod_{i=1}^N f(t_i) S(z_i) \text{ avec } f \text{ la densité de } Z$$

$$= \left[ \int \left( \int \mathbb{1}_{t \geq z} f(z) dz \right) f(t) dt \right]^{-N} \prod_{i=1}^N f(t_i) S(z_i) \text{ avec } \phi \text{ la FDR de } Z.$$

$\phi(t) = P(Z \leq t)$