# Modèles de durée discrets : Spécification et estimation

Note de travail

Bernard Lejeune
HEC - Université de Liège

Septembre 2005

# 1. Modèle de base

On commence par considérer le cas de base où les individus peuvent passer d'un état initial à un seul et unique autre état, complémentaire de l'état initial. De façon concrète, dans ce texte, on prendra comme exemple le passage de l'état "chômeur" à l'état "est sorti du chômage". Après avoir examiné ce cas de base, nous envisagerons le cas plus général où les individus peuvent passer d'un état initial à plusieurs autres états (destinations multiples) mutuellement exclusifs : par exemple, le passage de l'état "chômeur" aux états "a trouvé un travail à temps plein", "a trouvé un travail à temps partiel", "a quitté la population active", etc...

## 1.1. Variables et notations

- $T_i$ = durée de chômage, mesurée en mois, semaines ou jours, d'un individu $i$ pris au hasard dans une population donnée. Les valeurs possibles de $T_i$ sont supposées être $t = 0, 1, 2, ..., \infty$, $t = 0$ désignant la période d'entrée au chômage, $t = 1$ la première période suivant celle d'entrée au chômage, $t = 2$ la seconde période suivant celle d'entrée au chômage, etc...

- $X_{it}$ = vecteur de variables explicatives caractérisant l'individu $i$ et les conditions du marché du travail rencontrées par l'individu $i$ au cours de la période $t$ (= la $t^{ième}$ période suivant celle d'entrée au chômage).

- On cherche à modéliser la durée de chômage $T_i$ d'un individu $i$ pris au hasard au fonction des variables explicatives $X_{it}$ qui le caractérisent.

- Conventions de notation supplémentaires :

  a- $X^{it} = (X'_{i0}, X'_{i1}, ..., X'_{it})'$, vecteur de variables explicatives reprenant l'ensemble des variables caractérisant l'individu $i$ et les conditions du marché du travail rencontrées par l'individu $i$ au cours des périodes 0 à $t$.

  b- $X^{\underline{it+1}} = \left(X'_{it+1}, X'_{it+2}, ..., X'_{i\infty}\right)'$, vecteur de variables explicatives reprenant l'ensemble des variables caractérisant l'individu $i$ et les conditions du marché du travail rencontrées par l'individu $i$ au cours de toutes les périodes postérieures à $t$ : $t+1, t+2, ....$

  c- $X^{i\infty} = (X^{it\prime}, X^{\underline{it+1}\prime})' = (X'_{i0}, X'_{i1}, ..., X'_{i\infty})'$, vecteur de variables explicatives reprenant l'ensemble des variables caractérisant l'individu $i$ et les conditions du marché du travail rencontrées par l'individu $i$ au cours de toutes les périodes de $t = 0$ à $t = \infty$.

- Remarque :

  Si les variables explicatives $X_{it}$ ne varient pas dans le temps, on a $X_{it} = X_i$ $\forall\, t = 0, 1, 2, ...$, et $X^{it}, X^{\underline{it+1}}$ et $X^{i\infty}$ se réduisent tout simplement à $X_i$, où $X_i$ dénote le vecteur de variables explicatives invariantes dans temps caractérisant l'individu $i$.

## 1.2. Ingrédients du modèle

### 1.2.1. Fonction de hasard conditionnelle

- La fonction de hasard conditionnelle est définie par :

$$\lambda(t, X^{it}) = I\!P\left[T_i = t | T_i \geq t, X^{it}\right], \ \ \forall\, t = 0, 1, 2, ... \qquad (1)$$

  C'est la probabilité que la durée $T_i$ de chômage d'un individu $i$ pris au hasard soit (exactement) égale à $t$, sachant qu'elle est au moins égale à $t$ et les variables $X^{it}$. En termes de population, cette probabilité $\lambda(t, X^{it})$ s'interprète comme la proportion des individus qui sortent du chômage en (exactement) $t$ parmi les individus de la population qui sont toujours au chômage en $t-1$ et qui ont les caractéristiques décrites par $X^{it}$.

- La fonction de hasard conditionnelle $\lambda(t, X^{it})$ constitue le coeur du modèle. C'est la fonction que l'on cherche à estimer.

### 1.2.2. Exogénéité stricte des $X_{it}$

- On dit que les variables explicatives $X_{it}$ sont strictement exogènes si[1] :

[1] On notera que cette condition d'exogénéité peut être exprimée de façon équivalente et plus simple, mais

$$D\left(X^{\underline{it+1}}|T_i = t-k, X^{it}\right) = D\left(X^{\underline{it+1}}|X^{it}\right), \quad \begin{array}{l}\forall\, t = 0,1,2,...\\ \forall\, k = 0,1,2,...,t\end{array} \tag{2}$$

où $D\,(.|.)$ désigne les distributions conditionnelles de $X^{\underline{it+1}}$ sachant respectivement $(T_i = t-k, X^{it})$ et $X^{it}$, autrement dit si, sachant $X^{it}$, la connaissance que la durée $T_i$ est égale à $t-k$, $(\forall\, k = 0,1,2,...,t)$ n'apporte aucune information nouvelle pour prédire $X^{\underline{it+1}}$.

- Remarque :
  La condition d'exogénéité stricte de $X_{it}$ est trivialement satisfaite lorsque les variables explicatives $X_{it}$ sont invariantes dans le temps. On peut en pratique souvent, mais pas toujours, considérer cette condition d'exogénéité comme satisfaite.

- La condition (2) d'exogénéité stricte des $X_{it}$ implique (cf. preuve dans l'Annexe B) que :

$$I\!P\left[T_i = t|X^{it}, X^{\underline{it+1}}\right] = I\!P\left[T_i = t|X^{i\infty}\right] = I\!P\left[T_i = t|X^{it}\right], \quad \forall\, t = 0,1,2,... \tag{3}$$
$$I\!P\left[T_i \geq t|X^{it}, X^{\underline{it+1}}\right] = I\!P\left[T_i \geq t|X^{i\infty}\right] = I\!P\left[T_i \geq t|X^{it}\right], \quad \forall\, t = 0,1,2,... \tag{4}$$

et

$$\begin{aligned} I\!P\left[T_i = t|T_i \geq t, X^{it}, X^{\underline{it+1}}\right] &= I\!P\left[T_i = t|T_i \geq t, X^{i\infty}\right] \\ &= I\!P\left[T_i = t|T_i \geq t, X^{it}\right] \\ &= \lambda(t, X^{it}), \quad \forall\, t = 0,1,2,... \end{aligned} \tag{5}$$

autrement dit que, sachant $X^{it}$, la connaissance de $X^{\underline{it+1}}$ n'apporte aucune information sur la probabilité que la durée $T_i$ de chômage d'un individu $i$ pris au hasard soit égale à $t$, supérieure ou égale à $t$, ou encore égale à $t$ sachant qu'elle est supérieure ou égale à $t$.

### 1.2.3. Fonction de survie conditionnelle

- On appelle fonction de survie conditionnelle la quantité :

$$S(t, X^{it}) = I\!P\left[T_i \geq t|X^{it}\right], \quad \forall\, t = 0,1,2,...$$

C'est la probabilité que la durée $T_i$ de chômage d'un individu $i$ pris au hasard soit supérieur ou égale à $t$, sachant les variables $X^{it}$. En termes de population, cette probabilité $S(t, X^{it})$ s'interprète comme la proportion des individus qui sortent du chômage en $t$ ou après $t$, ou ce qui revient au même la proportion des individus qui sont toujours au chômage en $t-1$, parmi les individus de la population qui ont les caractéristiques décrites par $X^{it}$.

---

moins intuitive (cf. preuve dans l'Annexe A) par :

$$D\left(X^{\underline{it+1}}|T_i = t, X^{it}\right) = D\left(X^{\underline{it+1}}|X^{it}\right), \quad \forall\, t = 0,1,2,...$$

- Sous l'hypothèse d'exogénéité stricte des $X_{it}$, étant donné (4), on a :

$$S(t, X^{it}) = I\!P\left[T_i \geq t | X^{it}\right] = I\!P\left[T_i \geq t | X^{i\infty}\right] = \overline{S}(t|X^{i\infty}), \ \ \forall\, t = 0, 1, 2, ...$$

où la fonction $\overline{S}(t|X^{i\infty})$ est la fonction de survie[2] (conditionnelle) associée à la distribution conditionnelle définie par :

$$\overline{f}(t|X^{i\infty}) = I\!P\left[T_i = t | X^{i\infty}\right], \ \ \forall\, t = 0, 1, 2, ...$$

c'est-à-dire la distribution conditionnelle de $T_i$ sachant $X^{i\infty}$. Autrement dit, sous l'hypothèse d'exogénéité stricte des $X_{it}$, la fonction $S(t, X^{it})$ défini bien, comme son nom le suggère, une fonction de survie (conditionnelle) au sens habituel du terme, découlant d'une distribution conditionnelle bien définie.

- Si on renonce à l'hypothèse d'exogénéité stricte des $X_{it}$, contrairement à ce que son nom suggère, la fonction $S(t, X^{it})$ ne peut plus être interprétée comme une fonction de survie (conditionnelle) au sens habituel du terme, découlant d'une distribution conditionnelle bien définie. On pourrait être tenté de croire qu'il s'agit de la fonction de survie (conditionnelle) associée à la distribution conditionnelle qui serait définie par :

$$f(t, X^{it}) = I\!P\left[T_i = t | X^{it}\right], \ \ \forall\, t = 0, 1, 2, ...$$

C'est faux. En l'absence d'exogénéité stricte des $X_{it}$, la fonction $f(t, X^{it})$ ne définit en effet pas une distribution conditionnelle au sens habituel du terme (il faudrait pour cela que le vecteur de variables conditionnantes $X^{it}$ dépende pas de $t$), mais seulement une collection (pour $t = 0, 1, 2,...$) de probabilités conditionnelles. En bref, lorsqu'on renonce à l'hypothèse d'exogénéité stricte des $X_{it}$, la fonction $S(t, X^{it})$ ne définit pas une fonction de survie au sens habituel du terme.

- Sous l'hypothèse d'exogénéité stricte des $X_{it}$, de la fonction de hasard conditionnelle $\lambda(t, X^{it})$, on peut déduire la fonction de survie conditionnelle $S(t, X^{it})$. Etant donné (4) et (5) et que, de facon générale, $X^{it} = (X^{it-1\prime}, X_{it}^{\prime})^{\prime}$, où $X^{it-1} = \left(X_{i0}^{\prime}, X_{i1}^{\prime}, ..., X_{it-1}^{\prime}\right)^{\prime}$, on a[3] :

---

[2] On notera que cette définition (classique dans les modèles de durée) de la fonction de survie (conditionnelle) implique que son complément, c-à-d. la fonction de répartition (conditionnelle), est définie par $\overline{F}(t|X^{i\infty}) = 1 - \overline{S}(t|X^{i\infty}) = I\!P\left[T_i < t|X^{i\infty}\right]$, et non par la définition plus classique pour une v.a. discrète : $\overline{F}(t|X^{i\infty}) = I\!P\left[T_i \leq t|X^{i\infty}\right]$.

[3] La validité la relation de récurrence :

$$I\!P\left[T_i \geq t|X^{it}\right] = I\!P\left[T_i \geq t|T_i \geq t-1, X^{it}\right] I\!P\left[T_i \geq t-1|X^{it}\right]$$

vient du fait que l'événement $(T_i \geq t)$ implique l'événement $(T_i \geq t-1)$, et donc que l'intersection de ces deux événements est tout simplement l'événement $(T_i \geq t)$ :

$$I\!P\left[T_i \geq t, T_i \geq t-1|X^{it}\right] = I\!P\left[T_i \geq t|X^{it}\right]$$

$$\begin{aligned}
I\!P\left[T_i \geq 0|X^{i0}\right] &= 1 \\
I\!P\left[T_i \geq 1|X^{i1}\right] &= I\!P\left[T_i \geq 1|T_i \geq 0, X^{i1}\right] I\!P\left[T_i \geq 0|X^{i1}\right] \\
&= \left(1 - I\!P\left[T_i < 1|T_i \geq 0, X^{i1}\right]\right) I\!P\left[T_i \geq 0|X^{i1}\right] \\
&= \left(1 - I\!P\left[T_i = 0|T_i \geq 0, X^{i1}\right]\right) I\!P\left[T_i \geq 0|X^{i1}\right] \\
&= \left(1 - I\!P\left[T_i = 0|T_i \geq 0, X^{i0}\right]\right) I\!P\left[T_i \geq 0|X^{i0}\right] \\
&= \left(1 - \lambda(0, X^{i0})\right) \times 1 = 1 - \lambda(0, X^{i0}) \\
I\!P\left[T_i \geq 2|X^{i2}\right] &= I\!P\left[T_i \geq 2|T_i \geq 1, X^{i2}\right] I\!P\left[T_i \geq 1|X^{i2}\right] \\
&= \left(1 - I\!P\left[T_i < 2|T_i \geq 1, X^{i2}\right]\right) I\!P\left[T_i \geq 1|X^{i2}\right] \\
&= \left(1 - I\!P\left[T_i = 1|T_i \geq 1, X^{i2}\right]\right) I\!P\left[T_i \geq 1|X^{i2}\right] \\
&= \left(1 - I\!P\left[T_i = 1|, T_i \geq 1, X^{i1}\right]\right) I\!P\left[T_i \geq 1|X^{i1}\right] \\
&= \left(1 - \lambda(1, X^{i1})\right) \left(1 - \lambda(0, X^{i0})\right) \\
I\!P\left[T_i \geq 3|X^{i3}\right] &= I\!P\left[T_i \geq 3|T_i \geq 2, X^{i3}\right] I\!P\left[T_i \geq 2|X^{i3}\right] \\
&= \left(1 - I\!P\left[T_i < 3|T_i \geq 2, X^{i3}\right]\right) I\!P\left[T_i \geq 2|X^{i3}\right] \\
&= \left(1 - I\!P\left[T_i = 2|T_i \geq 2, X^{i3}\right]\right) I\!P\left[T_i \geq 2|X^{i3}\right] \\
&= \left(1 - I\!P\left[T_i = 2|, T_i \geq 2, X^{i2}\right]\right) I\!P\left[T_i \geq 2|X^{i2}\right] \\
&= \left(1 - \lambda(2, X^{i2})\right) \left(1 - \lambda(1, X^{i1})\right) \left(1 - \lambda(0, X^{i0})\right)
\end{aligned}$$

et ainsi de suite, de sorte que, de façon générale, on a :

$$S(t, X^{it}) = I\!P\left[T_i \geq t|X^{it}\right] = \begin{cases} 1 & \text{, si } t = 0 \\ \prod_{t^*=0}^{t-1} \left(1 - \lambda(t^*, X^{it^*})\right) & \text{, } \forall\, t = 1, 2, \ldots \end{cases}$$

- Remarque :

  On notera que dans la formule ci-dessus, $S(t, X^{it})$ ne dépend en fait que de $X^{it-1} = \left(X'_{i0}, X'_{i1}, \ldots, X'_{it-1}\right)'$. Autrement dit, sous l'hypothèse d'exogénéité stricte des $X_{it}$, on a en fait[4] :

$$\begin{aligned}
S(t, X^{it}) &= I\!P\left[T_i \geq t|X^{it}\right] = I\!P\left[T_i \geq t|X^{it-1}\right] \\
&= S(t, X^{it-1}) = \prod_{t^*=0}^{t-1} \left(1 - \lambda(t^*, X^{it^*})\right), \ \forall\, t = 1, 2, \ldots \qquad (6)
\end{aligned}$$

  Pour une autre façon, plus directe, de vérifier que, sous l'hypothèse d'exogénéité stricte des $X_{it}$, on a bien $I\!P[T_i \geq t|X^{it}] = I\!P[T_i \geq t|X^{it-1}]$, $\forall\, t = 1, 2, \ldots$, cf. Annexe C.

- Si on renonce l'hypothèse d'exogénéité stricte des $X_{it}$, on peut toujours exprimer la fonction $S(t, X^{it})$ par rapport à la fonction de hasard conditionnelle $\lambda(t, X^{it})$,

[4] On remarquera que cette égalité n'a aucun sens en $t = 0$.

mais elle n'a plus la forme simple obtenue ci-dessus. Elle prend la forme plus complexe (cf. preuve dans l'Annexe D) :

$$S(t,X^{it}) = I\!P\left[T_i \geq t|X^{it}\right] = \begin{cases} 1 & \text{, si } t=0 \\ \displaystyle\prod_{t^*=0}^{t-1}\left(1-\lambda(t^*,X^{it^*})\right) \\ \displaystyle\times \prod_{t^*=0}^{t-1} \frac{I\!P\left[X_{it^*+1}|T_i \geq t^*+1, X^{it^*}\right]}{I\!P\left[X_{it^*+1}|X^{it^*}\right]} & \text{, } \forall\, t=1,2,... \end{cases}$$

forme qui, comme on peut le voir, dépend notamment de la distribution des $X_{it}$, et qui, évidemment, contient comme cas particulier la formule développée ci-dessus sous l'hypothèse d'exogénéité stricte des $X_{it}$ (à nouveau, cf. preuve dans l'annexe D). En résumé, si on renonce à l'hypothèse d'exogénéité stricte des $X_{it}$, non seulement la fonction $S(t,X^{it})$ ne définit pas une fonction de survie au sens habituel du terme, et ne peut donc pas être interprétée comme telle, mais son expression en termes de la fonction de hasard conditionnelle $\lambda(t,X^{it})$ n'est plus aussi simple que dans le cas où les $X_{it}$ sont strictement exogènes et dépend notamment de la distribution des $X_{it}$.

### 1.2.4. Fonction de densité conditionnelle

- On appelle fonction de densité conditionnelle la quantité :

  $$f(t,X^{it}) = I\!P\left[T_i = t|X^{it}\right],\ \forall\, t=0,1,2,...$$

  C'est la probabilité que la durée $T_i$ de chômage d'un individu $i$ pris au hasard soit (exactement) égale à $t$, sachant les variables $X^{it}$. En termes de population, cette probabilité $f(t,X^{it})$ s'interprète comme la proportion des individus qui sortent du chômage en (exactement) $t$ parmi les individus de la population qui ont les caractéristiques décrites par $X^{it}$.

- Comme dans le cas de la fonction de survie conditionnelle $S(t,X^{it})$, sous l'hypothèse d'exogénéité stricte des $X_{it}$, étant donné (3), on a :

  $$f(t,X^{it}) = I\!P\left[T_i = t|X^{it}\right] = I\!P\left[T_i = t|X^{i\infty}\right] = \overline{f}(t|X^{i\infty}),\ \ \forall\, t=0,1,2,...$$

  où $\overline{f}(t|X^{i\infty})$ défini la distribution conditionnelle de $T_i$ sachant $X^{i\infty}$. Autrement dit, sous l'hypothèse d'exogénéité stricte des $X_{it}$, la fonction $f(t,X^{it})$ défini bien, comme son nom le suggère, une distribution (discrète) conditionnelle au sens habituel du terme.

- Comme déjà mentionné ci-dessus, si on renonce à l'hypothèse d'exogénéité stricte des $X_{it}$, contrairement à ce que son nom suggère, la fonction $f(t,X^{it})$ ne peut plus être interprétée comme une distribution (discrète) conditionnelle au sens habituel du terme : lorsqu'on renonce à l'hypothèse d'exogénéité stricte des $X_{it}$, la fonction $f(t,X^{it})$ ne définit effet pas une fonction de densité (discrète) au sens habituel du terme, mais seulement une collection (pour $t=0,1,2,...$) de probabilités conditionnelles.

- De la fonction de hasard conditionnelle $\lambda(t, X^{it})$, on peut à nouveau déduire la fonction de densité conditionnelle $f(t, X^{it})$. Par définition, de façon générale, on a[5] :

$$\begin{aligned} I\!P\left[T_i = t|X^{it}\right] &= I\!P\left[T_i = t|T_i \geq t, X^{it}\right] I\!P\left[T_i \geq t|X^{it}\right] \\ &= \lambda(t, X^{it})S(t, X^{it}), \quad \forall\, t = 0, 1, 2, ... \end{aligned}$$

de sorte que, en utilisant les expressions obtenues pour $S(t, X^{it})$ à la section précédente, sous l'hypothèse d'exogénéité stricte des $X_{it}$, on a :

$$f(t, X^{it}) = I\!P\left[T_i = t|X^{it}\right] = \begin{cases} \lambda(0, X^{i0}) & , \text{si } t = 0 \\ \lambda(t, X^{it}) \prod_{t^*=0}^{t-1} \left(1 - \lambda(t^*, X^{it^*})\right) & , \forall\, t = 1, 2, ... \end{cases}$$

tandis que si on renonce l'hypothèse d'exogénéité stricte des $X_{it}$, on a la forme plus complexe :

$$f(t, X^{it}) = I\!P\left[T_i = t|X^{it}\right] = \begin{cases} \lambda(0, X^{i0}) & , \text{si } t = 0 \\ \lambda(t, X^{it}) \prod_{t^*=0}^{t-1} \left(1 - \lambda(t^*, X^{it^*})\right) & \\ \times \prod_{t^*=0}^{t-1} \dfrac{I\!P\left[X_{it^*+1}|T_i \geq t^* + 1, X^{it^*}\right]}{I\!P\left[X_{it^*+1}|X^{it^*}\right]} & , \forall\, t = 1, 2, ... \end{cases}$$

forme qui, comme on peut le voir, dépend à nouveau notamment de la distribution des $X_{it}$. Ainsi, on a à nouveau que, si on renonce à l'hypothèse d'exogénéité stricte des $X_{it}$, non seulement la fonction $f(t, X^{it})$ ne définit pas une fonction de densité au sens habituel du terme, et ne peut donc pas être interprétée comme telle, mais encore que son expression en termes de la fonction de hasard conditionnelle $\lambda(t, X^{it})$ n'est plus aussi simple que dans le cas où les $X_{it}$ sont strictement exogènes et dépend notamment de la distribution des $X_{it}$.

## 1.3. Estimation du modèle

- On cherche à estimer la fonction de hasard conditionnelle $\lambda(t, X^{it})$ d'un individu $i$ tiré au hasard parmi les individus d'une population cible (par exemple, les individus qui s'inscrivent comme demandeur d'emploi au cours d'une période

[5] La validité de la relation :

$$I\!P\left[T_i = t|X^{it}\right] = I\!P\left[T_i = t|T_i \geq t, X^{it}\right] I\!P\left[T_i \geq t|X^{it}\right]$$

vient, de façon semblable au cas de la note 3, du fait que l'événement $(T_i = t)$ implique l'événement $(T_i \geq t)$, et donc que l'intersection de ces deux événements est tout simplement l'événement $(T_i = t)$ :

$$I\!P\left[T_i = t, T_i \geq t|X^{it}\right] = I\!P\left[T_i = t|X^{it}\right]$$

donnée).

- On commence par spécifier une forme fonctionnelle $\lambda(t, X^{it}; \beta)$, qui dépend d'un vecteur $p \times 1$ de paramètres $\beta \in \Theta$, pour la fonction de hasard conditionnelle $\lambda(t, X^{it})$, forme fonctionnelle que l'on suppose correctement spécifiée, i.e. telle que :
$$\lambda(t, X^{it}; \beta^o) = \lambda(t, X^{it}), \text{ pour une valeur } \beta^o \in \Theta$$

- La valeur la fonction de hasard conditionnelle doit toujours être comprise[6] entre 0 et 1. De façon à ce que cette restriction soit toujours satisfaite, on spécifiera par exemple :
$$\lambda(t, X^{it}; \beta) = g\left(X^{it*\prime}\beta\right)$$
où $g(x) = \dfrac{e^x}{1+e^x}$ (= fonction (cdf.) logistique) et $X^{it*}$ est un vecteur de variables composé de $t$, (d'un sous-ensemble)[7] de $X^{it}$, et de leur puissances (carré, cube, ...) et interactions (autrement dit, une forme polynomiale en $t$ et (le sous-ensemble choisi de) $X^{it}$). Bien que cela ne garantisse seulement que la positivité de la valeur de la fonction de hasard conditionnelle, une autre spécification a priori intéressante (pour l'interprétation des paramètres $\beta$ en terme de semi-élasticité) pour le fonction $g(x)$ est $g(x) = e^x$. Notons que lorsque le modèle discret considéré est regardé comme la version discrétisée d'un modèle en temps continu avec une fonction de hasard proportionnel donnée par $\lambda_c(t_c, X^{it_c}; \beta) = \lambda_o(t_c)\exp(X^{it_c'}\beta)$, alors la fonction fonction de hasard discrète correspondante est donnée comme ci-dessus par $\lambda(t, X^{it}; \beta) = g\left(X^{it*\prime}\beta\right)$, où $g(x) = 1 - e^{-e^x}$ (qui, comme la fonction logistique, est toujours comprise entre 0 et 1) et où $X^{it*}$ est composé de variables binaires pour chacune des périodes $t$ et des variables $X^{it}$, sans interactions entre $t$ et $X^{it}$.

- Contrairement aux fonctions de survie et de densité conditionnelles, la fonction de hasard conditionnelle a toujours un sens précis et interprétable d'un point de vue pratique, que l'hypothèse d'exogénéité stricte des $X_{it}$ soit ou non satisfaite. Que cette hypothèse soit ou non satisfaite, il est donc toujours légitime de chercher à estimer la fonction de hasard conditionnelle $\lambda(t, X^{it})$ d'un individu $i$ pris au hasard. Dans le cas où l'hypothèse d'exogénéité stricte des $X_{it}$ n'est pas satisfaite, on a simplement que, de la fonction de hasard conditionnelle, on ne peut pas déduire des fonctions de survie et de densité conditionnelles au sens habituel de ces termes. Dans la suite, on considère le problème de l'estimation des paramètres de la fonction de hasard conditionnelle $\lambda(t, X^{it}; \beta)$ supposée correctement spécifiée, tout d'abord sous l'hypothèse d'exogénéité stricte des $X_{it}$, et ensuite sans faire appel à cette hypothèse.

[6] La fonction de hasard conditionnelle donne une probabilité conditionnelle, et une probabilité conditionnelle est toujours comprise entre 0 et 1.

[7] En pratique, le vecteur de variables explicatives $X_{it}$ correspondant à la période $t$ considérée, dans lequel peut éventuellement être incorporé des variables explicatives retardées (avec un ou plusieurs lags).

### 1.3.1. Estimation sous l'hypothèse d'exogénéité stricte des $X_{it}$

- Si on suppose que les variables explicatives $X_{it}$ sont strictement exogènes, pour la forme fonctionnelle $\lambda(t, X^{it}; \beta)$ de la fonction de hasard conditionnelle, étant donné ce qui précède, on a comme forme fonctionnelle pour les fonctions de survie et de densité conditionnelle :

$$S(t, X^{it}; \beta) = \begin{cases} 1 & , \text{si } t = 0 \\ \prod_{t^*=0}^{t-1} \left(1 - \lambda(t^*, X^{it^*}; \beta)\right) & , \forall\, t = 1, 2, ... \end{cases} \tag{7}$$

et

$$f(t, X^{it}; \beta) = \lambda(t, X^{it}; \beta) S(t, X^{it}; \beta), \quad \forall\, t = 0, 1, 2, ...$$

- Sous l'hypothèse que la forme fonctionnelle $\lambda(t, X^{it}, \beta)$ est correctement spécifiée, on a évidemment que $S(t, X^{it}; \beta)$ et $f(t, X^{it}; \beta)$ sont correctement spécifiées, i.e. que :

$$\begin{cases} S(t, X^{it}; \beta^o) = S(t, X^{it}) \\ f(t, X^{it}; \beta^o) = f(t, X^{it}) \end{cases}, \quad \text{pour } \beta^o \in \Theta$$

- Ayant spécifié la fonction de hasard conditionnelle $\lambda(t, X^{it}; \beta)$ dont découlent, sous l'hypothèse d'exogénéité stricte des $X_{it}$, des spécifications pour les fonctions de survie et de densité conditionnelles, on peut, en supposant le modèle correctement spécifié et sur base d'un échantillon d'observations des durées $T_i$ de chômage et des variables explicatives $X_{it}$ de ces durées pour $n$ individus tirés au hasard, estimer la "vraie valeur" inconnue $\beta^o$ du vecteur de paramètres $\beta$ par la méthode du Maximum de Vraisemblance (MV).

- Pour rappel, de façon générique, s'appuyant sur l'observation d'un échantillon $((Y_1, X_1), ..., (Y_n, X_n))$ indépendamment (mais non nécessairement identiquement) distribué de réalisations d'un couple de variables aléatoires $(Y_i, X_i)$ — où tant $Y_i$ que $X_i$ peuvent être des vecteurs de variables — pour lequel est formulé un modèle paramétrique conditionnel :

$$\mathcal{P} \equiv \{f_i(y_i|X_i; \theta) : \theta \in \Theta \subset \mathbb{R}^p, i = 1, ..., n\}$$

  correctement spécifié, i.e. tel que :

$$f_i(y_i|X_i; \theta^o) = p_i\left(y_i|X_i\right), \quad \text{pour une valeur } \theta^o \in \Theta, \forall\, i = 1, ..., n \tag{8}$$

  où $p_i\left(y_i|X_i\right)$ représente la "vraie" densité conditionnelle de $Y_i$ sachant $X_i$, un estimateur MV de $\theta^o$ est donné par :

$$\hat{\theta}_n = \text{Argmax}_{\beta \in \Theta} \sum_{i=1}^{n} \ln f_i(Y_i|X_i; \theta) \tag{9}$$

  Sous des conditions de régularité générales, cet estimateur $\hat{\theta}_n$ est un estimateur convergent et asymptotiquement normal de $\theta^o$ :

$$\hat{\theta}_n \xrightarrow{p} \theta^o \quad \text{et} \quad V_n^{o\frac{1}{2}} \sqrt{n}(\hat{\theta}_n - \theta^o) \xrightarrow{d} \mathcal{N}(0, I_p)$$

où :

$$\begin{aligned} V_n^o &= \frac{1}{n}\sum_{i=1}^{n} E\left[\left(\frac{\partial \ln f_i(Y_i|X_i;\theta^o)}{\partial\beta}\right)\left(\frac{\partial \ln f_i(Y_i|X_i;\theta^o)}{\partial\beta}\right)'\right] \\ &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial^2 \ln f_i(Y_i|X_i;\theta^o)}{\partial\beta\partial\beta'}\right] \end{aligned}$$

soit en termes d'approximation utilisable est échantillon fini :

$$\hat{\theta}_n \approx \mathcal{N}(\theta^o, \hat{V}_n^{-1}/n)$$

où $\hat{V}_n$ est un estimateur convergent de $V_n^o$. En pratique, on utilise typiquement comme estimateur convergent de $V_n^o$ l'estimateur :

$$\hat{V}_n = -\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2 \ln f_i(Y_i|X_i;\hat{\theta}_n)}{\partial\beta\partial\beta'}$$

• Remarques :

a- On appelle un estimateur MV tel que défini ci-dessus, un estimateur Maximum de Vraisemblance *conditionnel* (MVc), car il s'appuye uniquement sur les densités conditionnelles de $Y_i$ sachant $X_i$, et non sur les densités jointes de $(Y_i, X_i)$.

b- L'estimateur $\hat{\theta}_n$ donné par (9) peut être exprimé de façon équivalente par :

$$\hat{\theta}_n = \text{Argmax}_{\beta\in\Theta} \prod_{i=1}^{n} f_i(Y_i|X_i;\theta)$$

et, étant l'hypothèse d'indépendance de l'ensemble des couples $\{(Y_i, X_i)\,,$ $i = 1, ..., n\}$, évalué à la "vraie valeur" $\theta^o$ de $\theta$, on a :

$$\prod_{i=1}^{n} f_i(y_i|X_i;\theta^o) = p\left(y_1, ..., y_n|X_1, ..., X_n\right)$$

où $p\left(y_i, ..., y_n|X_1, ..., X_n\right)$ représente la "vraie" densité conditionelle jointe de $(Y_1, ..., Y_n)$ sachant $(X_1, ..., X_n)$. Autrement dit, l'estimateur $\hat{\theta}_n$ de $\theta^o$ n'est autre que la valeur de $\theta$ qui maximise une forme paramétrique correctement spécifiée pour la densité conditionnelle *jointe* de $(Y_1, ..., Y_n)$ sachant $(X_1, ..., X_n)$ (qu'on appelle communément la (fonction de) *vraisemblance* (conditionnelle) de l'échantillon $((Y_1, X_1)\,, ..., (Y_n, X_n))$, la *log-vraisemblance* (conditionelle) de l'échantillon étant simplement le logarithme (naturel) de cette quantité : $\ln\left(\prod_{i=1}^{n} f_i(y_i|X_i;\theta)\right) = \sum_{i=1}^{n} \ln f_i(y_i|X_i;\theta)$).

c- Pour que l'estimateur $\hat{\theta}_n$ défini par (9) soit un estimateur convergent et asymptotiquement normal de $\theta^o$, il n'est pas nécessaire que l'ensemble des couples $\{(Y_i, X_i)\ ,\ i = 1, ..., n\}$ de l'échantillon soient indépendants, ou, condition moins restrictive[8], que le produit $\prod_{i=1}^{n} f_i(y_i|X_i; \theta^o)$ des densités conditionnelles spécifiées corresponde bien, évalué à la "vraie valeur" $\theta^o$ de $\theta$, à la "vraie" densité conditionnelle jointe de $(Y_1, ..., Y_n)$ sachant $(X_1, ..., X_n)$ : il suffit que le modèle paramétrique conditionnel $\mathcal{P}$ soit correctement spécifié pour les densités conditionnelles ("marginales") de $Y_i$ sachant $X_i$, $\forall\, i = 1, ..., n$ (= la condition (8))[9].

De façon générale, un estimateur MV définit comme la valeur d'un paramètre qui maximise le produit (ou de façon équivalente la somme des logarithmes (naturels)) d'un ensemble de densités conditionnelles (pas nécessairement de type "marginales" comme ci-dessus) correctement spécifiées, et tel que le produit de ces densités conditionnelles n'est pas, évalué à la "vraie valeur" du paramètre, égal à la "vraie" densité conditionnelle jointe des observations l'échantillon, est appelé un estimateur Maximum de Vraisemblance conditionnel *partiel* (MVcp). De tels estimateurs sont utilisés de façon extensive en économétrie. Ils possèdent des propriétés semblables à celles des estimateurs MVc standards : convergence et normalité asymptotique, avec néanmoins comme différence notable une expression généralement (mais pas toujours ; dans beaucoup de cas courants, elles sont identiques) plus complexe que celle des estimateurs MVc standards pour la matrice de variance-covariance asymptotique de l'estimateur.

- La forme de l'estimateur MV (ou plus précisément MVc) de la "vraie valeur" inconnue $\beta^o$ du paramètre $\beta$ de la fonction de hasard conditionnel $\lambda(t, X^{it}; \beta)$ dépend du schéma d'échantillonnage sur base duquel on suppose qu'est obtenu un échantillon d'observations sur les durées $T_i$ de chômage et les variables explicatives $X_{it}$ de ces durées, et en particulier du fait de savoir si ces durées et ces variables explicatives sont observées de façon complète ou non (la cas typique est le cas où les durées et les variables explicatives ne sont pas observées de façon complète). On considère ci-dessous trois schémas d'échantillonnage, par ordre croissant de complexité : échantillonnage aléatoire simple, échantillonnage aléatoire avec censure à droite, et finalement échantillonnage aléatoire avec censure à droite et observation (partielle) par intervalles. Ces trois schémas d'échantillonnage correspondent à trois cas de figure différents concernant l'observabilité des durées $T_i$ de chômage et des variables explicatives $X_{it}$ de ces durées. Pour chacun d'eux , on dérive l'estimateur MVc de $\beta^o$. On notera le cas de l'échantillonnage aléatoire simple peut être regardé comme un cas particulier de l'échantillonnage aléatoire avec censure à droite, qui lui-même peut être regardé comme un cas particulier de l'échantillonnage aléatoire avec censure à droite et observation (partielle) par intervalles En d'autres termes, ce dernier cas couvre en fait les deux cas précédents.

---

[8] Qui suppose que les $\{Y_i,\ i = 1, ..., n\}$ sont indépendants conditionnellement à $(X_1, ..., X_n)$. On notera que cette condition est suffisante pour toujours avoir la même expression simple que celle décrite ci-dessus pour la matrice de variance-covariance asymptotique de l'estimateur.

[9] On notera cependant que, dans ce cas de figure, la matrice de variance-covariance asymptotique de l'estimateur n'a généralement plus la forme simple décrite ci-dessus.

**A. Echantillonnage aléatoire simple**

- Dans ce schéma d'échantillonnage, on suppose qu'un échantillon d'observations de taille $n$ est obtenu en tirant au hasard $n$ individus dans la population cible, et que pour chaque individu tiré, on observe de façon complète la durée $T_i$ de son épisode de chômage, ainsi que l'ensemble des variables explicatives $X_{it}$ de $t = 0$ jusque $t = T_i$. Pour chaque individu tiré, on observe donc une réalisation du couple $(T_i, X^{iT_i})$, et l'échantillon d'observations pour les $n$ individus tirés est une réalisation du $n$-uple de couples $((T_1, X^{1T_1}), ..., (T_n, X^{nT_n}))$.

  On notera que ce mode d'échantillonnage suppose que chacun des individus tirés soit suivi jusqu'à la fin de son épisode de chômage, qu'elle qu'en soit la longueur. Il s'agit évidemment d'un mode d'échantillonnage peu probable en pratique, mais qu'il est intéressant de considérer à titre de référence.

- On commence par raisonner comme si, pour chaque individu tiré, on observait une réalisation du couple $(T_i, X^{i\infty})$, i.e. comme si les variables explicatives $X_{it}$ étaient toujours entièrement observées, de $t = 0$ jusque $t = \infty$. Dans ce cas de figure, étant donné l'hypothèse d'exogénéité stricte des $X_{it}$, si le modèle est bien correctement spécifié, la distribution conditionnelle à $X^{i\infty}$ de la durée $T_i$ de chômage d'un individu $i$ tiré au hasard est donnée par la fonction de densité conditionnelle :

$$\begin{aligned} \overline{f}(t|X^{i\infty}) &= I\!P\left[T_i = t|X^{i\infty}\right] = I\!P\left[T_i = t|X^{it}\right] \\ &= f(t, X^{it}; \beta^o) = \lambda(t, X^{it}; \beta^o) S(t, X^{it}; \beta^o), \quad \forall\, t = 0, 1, 2, ... \end{aligned}$$

  Pour un échantillon de $n$ individus tirés au hasard[10], les $n$ couples $(T_1, X^{1\infty})$, ..., $(T_n, X^{n\infty})$ sont indépendants et de même distribution, et un estimateur MVc $\hat{\beta}_n$ de $\beta^o$ est dès lors donné par :

$$\begin{aligned} \hat{\beta}_n &= \text{Argmax}_{\beta\in\Theta} \sum_{i=1}^{n} \ln f(T_i, X^{iT_i}; \beta) \\ &= \text{Argmax}_{\beta\in\Theta} \sum_{i=1}^{n} \ln \lambda(T_i, X^{iT_i}; \beta) + \sum_{i=1}^{n} \ln S(T_i, X^{iT_i}; \beta) \end{aligned}$$

  où :

$$\ln S(T_i, X^{iT_i}; \beta) = \begin{cases} 0 & , \text{si } T_i = 0 \\ \displaystyle\sum_{t^*=0}^{T_i - 1} \ln\left(1 - \lambda(t^*, X^{it^*}; \beta)\right) & , \forall\, T_i = 1, 2, ... \end{cases}$$

- On constate que l'estimateur MVc $\hat{\beta}_n$ de $\beta^o$ ne dépend pas de l'ensemble des variables de l'échantillon $((T_1, X^{1\infty}), ..., (T_n, X^{n\infty}))$, mais, étant donné l'hypothèse d'exogénéité stricte des $X_{it}$, seulement de $((T_1, X^{1T_1}), ..., (T_n, X^{nT_n}))$. Autrement dit, pour estimer $\beta^o$ par l'estimateur MVc $\hat{\beta}_n$, il n'est pas nécessaire de disposer d'observations de $((T_1, X^{1\infty}), ..., (T_n, X^{n\infty}))$, mais seulement d'observations de

[10] Au sens strict, avec remise.

$\left((T_1, X^{1T_1}\right), ..., (T_n, X^{nT_n}))$. $\hat{\beta}_n$ est donc un estimateur MVc adéquat pour estimer $\beta^o$ sur base des observations dont on dispose dans le schéma d'échantillonnage supposé.

- L'estimateur $\hat{\beta}_n$ étant un estimateur MVc standard, sous des conditions de régularité générales, il possède les propriétés habituelles des estimateurs MVc, à savoir (si, comme supposé, le modèle est bien correctement spécifié et les $X_{it}$ strictement exogènes) :

$$\hat{\beta}_n \xrightarrow{p} \beta^o \quad \text{et} \quad V^{o\frac{1}{2}}\sqrt{n}(\hat{\beta}_n - \beta^o) \xrightarrow{d} \mathcal{N}(0, I_p)$$

où[11] :

$$\begin{aligned} V^o &= \frac{1}{n}\sum_{i=1}^{n} E\left[\left(\frac{\partial \ln f(T_i, X^{iT_i}; \beta^o)}{\partial \beta}\right)\left(\frac{\partial \ln f(T_i, X^{iT_i}; \beta^o)}{\partial \beta}\right)'\right] \\ &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial^2 \ln f(T_i, X^{iT_i}; \beta^o)}{\partial \beta \partial \beta'}\right] \end{aligned}$$

soit en termes d'approximation utilisable est échantillon fini :

$$\hat{\beta}_n \approx \mathcal{N}(\beta^o, \hat{V}_n^{-1}/n)$$

avec

$$\hat{V}_n = -\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2 \ln f(T_i, X^{iT_i}; \hat{\beta}_n)}{\partial \beta \partial \beta'}$$

### B. Echantillonnage aléatoire avec censure à droite

- Dans ce schéma d'échantillonnage, on suppose qu'un échantillon de taille $n$ est obtenu en tirant au hasard $n$ individus dans la population cible, et que chaque individu tiré est observé, tant en ce qui concerne la durée $T_i$ de son épisode de chômage qu'en ce qui concerne l'ensemble de ses variables explicatives $X_{it}$, au maximum seulement durant une période allant de $t = 0$ jusque $t = C_i$, où $C_i = 0, 1, 2, ...$ désigne un point de censure à droite, fixe ou aléatoire.

  Sont ainsi observés, lorsque la durée $T_i$ de l'épisode de chômage est inférieur ou égale à $C_i$, à la fois la durée $T_i$ exacte de l'épisode de chômage et l'ensemble des variables explicatives $X_{it}$ de $t = 0$ jusque $t = T_i$, et lorsque la durée $T_i$ de l'épisode de chômage est supérieure ou égale à $C_i + 1$ (= supérieure à $C_i$), simplement le fait que l'on a $T_i \geq C_i + 1$ et l'ensemble des variables explicatives $X_{it}$ de $t = 0$ jusque $t = C_i$ (et non $C_i + 1$).

  Pour chaque individu tiré, on observe donc plus, comme dans le cas de l'échantil-

[11] On notera que la matrice de variance-covariance $V^o$ ne dépend pas de $n$. Cela résulte simplement du fait que, comme l'échantillon $((T_1, X^{1\infty}), ..., (T_n, X^{n\infty}))$ est i.i.d., on a :

$$E\left[\left(\frac{\partial \ln f(T_i, X^{iT_i}; \beta^o)}{\partial \beta}\right)\left(\frac{\partial \ln f(T_i, X^{iT_i}; \beta^o)}{\partial \beta}\right)'\right] = -E\left[\frac{\partial^2 \ln f(T_i, X^{iT_i}; \beta^o)}{\partial \beta \partial \beta'}\right] = V^o, \quad \forall\, i = 1, ..., n$$

lonnage aléatoire simple, une réalisation du couple $(T_i, X^{iT_i})$, mais une réalisation du triplet $(T_i^*, X^{iM_i^*}, C_i)$, où :

- $T_i^*$ désigne la durée de chômage effectivement observée de l'individu $i$ étant donné le mécanisme d'échantillonnage, durée qui est égale à :

$$T_i^* = \begin{cases} T_i & , \text{si } T_i \leq C_i \\ C_i + 1 & , \text{si } T_i \geq C_i + 1 \end{cases}$$

  où $T_i^*$ est par convention fixé à $C_i + 1$ lorsque $T_i \geq C_i + 1$ ($\Leftrightarrow T_i > C_i$), de sorte que l'ensemble $\mathbb{T}_i$ des valeurs possibles de $T_i^*$ est $\mathbb{T}_i \equiv \{0, 1, 2, ..., C_i, C_i + 1\}$.
- $M_i^* = \min(T_i^*, C_i)$, ce qui reflète le fait que, lorsque $T_i \leq C_i$ (et donc $T_i^* = T_i$), on observe les variables explicatives $X_{it}$ de $t = 0$ jusque $t = T_i^* = T_i$, et lorsque $T_i \geq C_i + 1$ (et donc $T_i^* = C_i + 1$), on observe les variables explicatives $X_{it}$ de $t = 0$ jusque $t = T_i^* - 1 = C_i$ (et non $T_i^* = C_i + 1$).

et l'échantillon d'observations pour les $n$ individus tirés est une réalisation du $n$-uple de triplets $((T_1^*, X^{1M_1^*}, C_1), ..., (T_n^*, X^{nM_n^*}, C_n))$.

Ce mode d'échantillonnage est le mode d'échantillonnage typiquement rencontré en pratique lorsqu'on travaille avec des données de durée.

- On raisonne dans un premier temps comme si, pour chaque individu tiré, on observait une réalisation du triplet $(T_i^*, X^{i\infty}, C_i)$, i.e. comme si les variables explicatives $X_{it}$ étaient toujours entièrement observées, de $t = 0$ jusque $t = \infty$.

  On commence par supposer que, conditionnellement à $X^{i\infty}$, la durée (qu'elle soit (si $T_i \leq C_i$) ou non (si $T_i \geq C_i + 1$) effectivement observée de façon exacte) de chômage $T_i$ d'un individu $i$ pris au hasard est indépendante de son point de censure $C_i$, soit que l'on a :

$$I\!P\left[T_i = t | X^{i\infty}, C_i\right] = I\!P\left[T_i = t | X^{i\infty}\right], \quad \forall\, t = 0, 1, 2, ... \tag{10}$$

  autrement dit que, sachant $X^{i\infty}$, la connaissance du point de censure $C_i$, c'est-à-dire que l'on observe de façon complète le parcours de l'individu $i$ seulement durant la période allant de $t = 0$ jusque $t = C_i$, n'apporte aucune information sur sa durée effective (simplement non observée de façon exacte si $T_i \geq C_i + 1$) de chômage $T_i$. Cette hypothèse est automatiquement satisfaite si le point de censure $C_i$ est fixe, non aléatoire, ou encore si c'est une fonction de $X^{i\infty}$.

  La condition (10) d'indépendance conditionnelle entre $T_i$ et $C_i$ implique (cf. preuve dans l'Annexe E) que :

$$I\!P\left[T_i \geq t | X^{i\infty}, C_i\right] = I\!P\left[T_i \geq t | X^{i\infty}\right], \quad \forall\, t = 0, 1, 2, ... \tag{11}$$

et par conséquent aussi que[12] :

$$I\!P\left[T_i = t | T_i \geq t, X^{i\infty}, C_i\right] = I\!P\left[T_i = t | T_i \geq t, X^{i\infty}\right], \quad \forall\, t = 0, 1, 2, ...$$

Etant donné la condition (10) d'indépendance conditionnelle entre $T_i$ et $C_i$, et son implication (11), on peut déduire la distribution conditionnelle à $(X^{i\infty}, C_i)$ de la durée de chômage effectivement observée $T_i^*$ d'un individu $i$ pris au hasard. Elle est donnée par la fonction de densité conditionnelle :

$$\overline{f}^*(t|X^{i\infty}, C_i) = I\!P\left[T_i^* = t|X^{i\infty}, C_i\right], \qquad \forall\, t = 0, 1, 2, ..., C_i + 1$$

$$= \begin{cases} I\!P[T_i = t|X^{i\infty}, C_i] = I\!P[T_i = t|X^{i\infty}] & \text{, si } t \leq C_i \\ I\!P[T_i \geq t|X^{i\infty}, C_i] = I\!P[T_i \geq t|X^{i\infty}] & \text{, si } t = C_i + 1 \end{cases}$$

qui, puisqu'étant donné l'hypothèse d'exogénéité stricte des $X_{it}$, si le modèle est bien correctement spécifié, on a :

$$\begin{aligned} I\!P\left[T_i = t|X^{i\infty}\right] &= I\!P\left[T_i = t|X^{it}\right] = f(t, X^{it}; \beta^o) \\ &= \lambda(t, X^{it}; \beta^o) S(t, X^{it}; \beta^o), \;\; \forall\, t = 0, 1, 2, ... \end{aligned}$$

et

$$\begin{aligned} I\!P\left[T_i \geq t|X^{i\infty}\right] &= I\!P\left[T_i \geq t|X^{it}\right] \\ &= S(t, X^{it}; \beta^o), \quad \forall\, t = 0, 1, 2, ... \end{aligned}$$

et que, étant donné (6)[13], on a également :

$$\begin{aligned} S(t, X^{it}; \beta^o) &= I\!P\left[T_i \geq t|X^{it}\right] = I\!P\left[T_i \geq t|X^{it-1}\right] \\ &= S(t, X^{it-1}; \beta^o) = \prod_{t^*=0}^{t-1} \left(1 - \lambda(t^*, X^{it^*}; \beta^o)\right), \quad \forall\, t = 1, 2, ... \end{aligned} \tag{12}$$

est égale à :

$$\begin{aligned} &\overline{f}^*(t|X^{i\infty}, C_i) = I\!P\left[T_i^* = t|X^{i\infty}, C_i\right] \\ = \; &I\!P\left[T_i^* = t|X^{im_i^*}, C_i\right], \;\; \forall\, t = 0, 1, 2, ..., C_i + 1 \\ = \; &f^*(t, X^{im_i^*}, C_i; \beta^o) = \begin{cases} \lambda(t, X^{it}; \beta^o) S(t, X^{it}; \beta^o) & \text{, si } t \leq C_i \\ S(t, X^{it-1}; \beta^o) & \text{, si } t = C_i + 1 \end{cases} \end{aligned}$$

où $m_i^* = \min(t, C_i)$.

Pour un échantillon de $n$ individus tirés au hasard[14], les $n$ triplets $(T_1^*, X^{1\infty}, C_1)$, ..., $(T_n^*, X^{n\infty}, C_n)$ sont indépendants et de même distribution, et un estimateur

---

[12] En effet, étant donné (10) et (11) et le fait que $(T_i = t) \Rightarrow (T_i \geq t)$, $\forall\, t = 0, 1, 2, ...$, on a :

$$I\!P\left[T_i = t|T_i \geq t, X^{i\infty}, C_i\right] = \frac{I\!P[T_i = t|X^{i\infty}, C_i]}{I\!P[T_i \geq t|X^{i\infty}, C_i]} = \frac{I\!P[T_i = t|X^{i\infty}]}{I\!P[T_i \geq t|X^{i\infty}]} = I\!P\left[T_i = t|T_i \geq t, X^{i\infty}\right]$$

[13] C'est-à-dire que la fonction de survie $S(t, X^{it})$ ne dépend en fait que de $X^{it-1}$.

[14] Au sens strict, avec remise.

MVc $\hat{\beta}_n$ de $\beta^o$ est dès lors donné par :

$$\begin{aligned}
\hat{\beta}_n &= \text{Argmax}_{\beta\in\Theta}\sum_{i=1}^{n}\ln f^*(T_i^*, X^{iM_i^*}, C_i;\beta) \\
&= \text{Argmax}_{\beta\in\Theta}\sum_{i=1}^{n}\Big[\,\mathbb{I}_{(T_i^*\leq C_i)}\big(\ln\lambda(T_i^*, X^{iT_i^*},\beta)+\ln S(T_i^*, X^{iT_i^*},\beta)\big) \\
&\qquad\qquad + \mathbb{I}_{(T_i^*=C_i+1)}\ln S(T_i^*, X^{iT_i^*-1},\beta)\,\Big]
\end{aligned}$$

qu'on peut encore écrire :

$$\begin{aligned}
\hat{\beta}_n &= \text{Argmax}_{\beta\in\Theta}\sum_{i=1}^{n}\ln f^*(T_i^*, X^{iM_i^*}, C_i;\beta) \\
&= \text{Argmax}_{\beta\in\Theta}\sum_{i=1}^{n}\mathbb{I}_{(T_i^*\leq C_i)}\ln\lambda(T_i^*, X^{iT_i^*},\beta)+\sum_{i=1}^{n}\ln S(T_i^*, X^{iM_i^*},\beta)
\end{aligned}$$

où $M_i^* = \min(T_i^*, C_i)$, les $\mathbb{I}_{(.)}$ sont des variables indicatrices et :

$$\ln S(T_i^*, X^{iM_i^*},\beta) = \begin{cases} 0 & , \text{ si } T_i^* = 0 \\ \displaystyle\sum_{t^*=0}^{T_i^*-1}\ln\big(1-\lambda(t^*, X^{it^*};\beta)\big) & , \forall\, T_i^* = 1, 2, ..., C_i+1 \end{cases}$$

- De façon semblable au cas de l'échantillonnage aléatoire simple, on constate que l'estimateur MVc $\hat{\beta}_n$ de $\beta^o$ ne dépend pas de l'ensemble des variables de l'échantillon $((T_1^*, X^{1\infty}, C_1), ..., (T_n^*, X^{n\infty}, C_n))$, mais, étant donné l'hypothèse d'exogénéité stricte des $X_{it}$ (et (12)), seulement de $((T_1^*, X^{1M_1^*}, C_1), ..., (T_n^*, X^{nM_n^*}, C_n))$. Autrement dit, pour estimer $\beta^o$ par l'estimateur MVc $\hat{\beta}_n$, il n'est pas nécessaire de disposer d'observations de $((T_1^*, X^{1\infty}, C_1), ..., (T_n^*, X^{n\infty}, C_n))$, mais seulement d'observations de $((T_1^*, X^{1M_1^*}, C_1), ..., (T_n^*, X^{nM_n^*}, C_n))$. $\hat{\beta}_n$ est donc un estimateur MVc adéquat pour estimer $\beta^o$ sur base des observations dont on dispose dans le schéma d'échantillonnage supposé.

- L'estimateur $\hat{\beta}_n$ étant un estimateur MVc standard, sous des conditions de régularité générales, il possède les propriétés habituelles des estimateurs MVc, à savoir (si, comme supposé, le modèle est bien correctement spécifié et les $X_{it}$ strictement exogènes) :

$$\hat{\beta}_n \xrightarrow{p} \beta^o \quad \text{et} \quad V^{o\frac{1}{2}}\sqrt{n}(\hat{\beta}_n-\beta^o)\xrightarrow{d}\mathcal{N}(0, I_p)$$

où[15] :

$$\begin{aligned}V^o &= \frac{1}{n}\sum_{i=1}^{n} E\left[\left(\frac{\partial \ln f^*(T_i^*, X^{iM_i^*}, C_i; \beta^o)}{\partial \beta}\right)\left(\frac{\partial \ln f^*(T_i^*, X^{iM_i^*}, C_i; \beta^o)}{\partial \beta}\right)'\right] \\ &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial^2 \ln f^*(T_i^*, X^{iM_i^*}, C_i; \beta^o)}{\partial \beta \partial \beta'}\right]\end{aligned}$$

soit en termes d'approximation utilisable est échantillon fini :

$$\hat{\beta}_n \approx \mathcal{N}(\beta^o, \hat{V}_n^{-1}/n)$$

avec

$$\hat{V}_n = -\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2 \ln f^*(T_i^*, X^{iM_i^*}, C_i; \hat{\beta}_n)}{\partial \beta \partial \beta'}$$

### C. Echantillonnage aléatoire avec censure à droite et observation (partielle) par intervalles

- Dans ce dernier schéma d'échantillonnage, on suppose, comme dans le cas précédent, qu'un échantillon de taille $n$ est obtenu en tirant au hasard $n$ individus dans la population cible, et que chaque individu tiré est observé, tant en ce qui concerne la durée $T_i$ de son épisode de chômage qu'en ce qui concerne l'ensemble de ses variables explicatives $X_{it}$, au maximum seulement durant une période allant de $t = 0$ jusque $t = C_i$, où $C_i$ désigne un point de censure à droite, fixe ou aléatoire.

  On ne suppose cependant plus que, durant la période allant de $t = 0$ jusque $t = C_i$, la durée $T_i$ de son épisode de chômage est observée de façon complète, mais bien, d'une part, observée par intervalles sur la période allant de $t = 0$ jusque $t = B_{iH}$ (avec $B_{iH} < C_i$), les $H$ intervalles d'observations ($H$ fixe) étant donnés par les intervalles $[0, B_{i1}]$ , $[B_{i1} + 1, B_{i2}], [B_{i2} + 1, B_{i3}], ..., [B_{ih-1} + 1, B_{ih}], ..., [B_{iH-1} + 1, B_{iH}]$, où $B_{i1}, B_{i2}, ..., B_{ih}, ..., B_{iH-1}, B_{iH}$ (avec $B_{i1} < B_{i2} < ... < B_{ih} < ... < B_{iH-1} < B_{iH}$) désignent les $H$ bornes supérieures des intervalles, fixes ou aléatoires, et d'autre part, observée de façon complète sur la période allant de $t = B_{iH} + 1$ jusque $t = C_i$, les variables explicatives $X_{it}$ étant elles supposées toujours observées de $t = 0$ jusque la fin $t = B_{ih}$ de l'intervalle lorsque la durée $T_i$ tombe dans l'intervalle $[B_{ih-1} + 1, B_{ih}]$ ($\forall\, h = 1, 2, ..., H$ , en posant par convention $B_{i0} = -1$), et de $t = 0$ jusque $t = T_i$ lorsque la durée $T_i$ est comprise (bornes incluses) entre $B_{iH} + 1$ et $C_i$, i.e. lorsque la durée $T_i$ est observée de façon complète.

---

[15] A nouveau, on notera que la matrice de variance-covariance $V^o$ ne dépend pas de $n$. Cela résulte simplement du fait que, comme l'échantillon $((T_1^*, X^{1\infty}, C_1), ..., (T_n^*, X^{n\infty}, C_n))$ est i.i.d., on a :

$$\begin{aligned}& E\left[\left(\frac{\partial \ln f^*(T_i^*, X^{iM_i^*}, C_i; \beta^o)}{\partial \beta}\right)\left(\frac{\partial \ln f^*(T_i^*, X^{iM_i^*}, C_i; \beta^o)}{\partial \beta}\right)'\right] \\ &= -E\left[\frac{\partial^2 \ln f^*(T_i^*, X^{iM_i^*}, C_i; \beta^o)}{\partial \beta \partial \beta'}\right] = V^o, \;\; \forall\, i = 1, ..., n\end{aligned}$$

On suppose par contre que rien ne change lorsque la durée $T_i$ de l'épisode de chômage est supérieure ou égale à $C_i + 1$ (= supérieure à $C_i$) : est dans ce cas simplement observé le fait que l'on a $T_i \geq C_i + 1$ et l'ensemble des variables explicatives $X_{it}$ de $t = 0$ jusque $t = C_i$ (et non $C_i + 1$).

Sont ainsi observés, lorsque la durée $T_i$ de l'épisode de chômage tombent dans le $h^{ème}$ intervalle $[B_{ih-1} + 1, B_{ih}]$, simplement le fait que $B_{ih-1} + 1 \leq T_i \leq B_{ih}$ et l'ensemble des variables explicatives $X_{it}$ de $t = 0$ jusque $t = B_{ih}$, lorsque la durée $T_i$ est comprise (bornes incluses) entre $B_{iH} + 1$ et $C_i$, à la fois la durée $T_i$ exacte de l'épisode de chômage et l'ensemble des variables explicatives $X_{it}$ de $t = 0$ jusque $t = T_i$, et enfin lorsque la durée $T_i$ de l'épisode de chômage est supérieure ou égale à $C_i + 1$ (= supérieure à $C_i$), simplement le fait que l'on a $T_i \geq C_i + 1$ et l'ensemble des variables explicatives $X_{it}$ de $t = 0$ jusque $t = C_i$ (et non $C_i + 1$).

Selon ce schéma d'échantillonage, pour chaque individu tiré, on observe donc une réalisation du $(H + 3)$-uple de variables aléatoires $(T_i^*, X^{iM_i^*}, B_{i1}, ..., B_{iH}, C_i)$, avec par hypothèse toujours $B_{i1} < B_{i2} < ... < B_{iH} < C_i$, où :

- $T_i^*$ désigne la durée de chômage effectivement observée de l'individu $i$ étant donné le mécanisme d'échantillonnage, durée qui est égale à :

$$T_i^* = \begin{cases} B_{ih} & \text{, si } B_{ih-1} + 1 \leq T_i \leq B_{ih} \\ T_i & \text{, si } B_{iH} + 1 \leq T_i \leq C_i \\ C_i + 1 & \text{, si } T_i \geq C_i + 1 \end{cases}$$

  où $T_i^*$ est par convention fixé à $B_{ih}$ lorsque $T_i$ tombe dans le $h^{ème}$ intervalle $[B_{ih-1} + 1, B_{ih}]$ et à $C_i + 1$ lorsque $T_i \geq C_i + 1$ ($\Leftrightarrow T_i > C_i$), de sorte que l'ensemble $\mathbb{T}_i$ des valeurs possibles de $T_i^*$ est $\mathbb{T}_i \equiv \{B_{i1}, ..., B_{iH}, B_{iH} + 1, ..., C_i, C_i + 1\}$, et par convention également $B_{i0} = -1$.
- $M_i^* = \min(T_i^*, C_i)$, ce qui reflète le fait que, lorsque $T_i \leq C_i$ (et donc $T_i^* = B_{ih}$ ou $T_i^* = T_i$), on observe les variables explicatives $X_{it}$ de $t = 0$ jusque $t = T_i^*$ (avec, selon les cas, $T_i^* = B_{ih}$ ou $T_i^* = T_i$) et lorsque $T_i \geq C_i + 1$ (et donc $T_i^* = C_i + 1$), on observe les variables explicatives $X_{it}$ de $t = 0$ jusque $t = T_i^* - 1 = C_i$ (et non $T_i^* = C_i + 1$).

et l'échantillon d'observations pour les $n$ individus tirés est une réalisation du $n$-uple de $(H + 3)$-uples $((T_1^*, X^{1M_1^*}, B_{11}, ..., B_{1H}, C_1), ...., (T_n^*, X^{nM_n^*}, B_{n1}, ..., B_{nH}, C_n))$.

Comme déjà évoqué, le présent schéma d'échantillonnage contient comme cas particulier l'échantillonnage aléatoire avec censure à droite examiné à la section précédente : il s'y réduit lorsque, pour $H$ quelconque, $B_{i1} = 0$, $B_{i2} = 1$, $B_{i3} = 2$, ..., $B_{iH} = H - 1$ (et on a toujours $C_i > B_{iH} = H - 1$).

- On raisonne dans un premier temps comme si, pour chaque individu tiré, on observait une réalisation du $(H + 3)$-uple $(T_i^*, X^{i\infty}, B_{i1}, ..., B_{iH}, C_i)$, i.e. comme si les variables explicatives $X_{it}$ étaient toujours entièrement observées, de $t = 0$

jusque $t = \infty$.

On commence par supposer que, conditionnellement à $X^{i\infty}$, la durée (qu'elle soit ou non effectivement observée de façon exacte) de chômage $T_i$ d'un individu $i$ pris au hasard est indépendante des bornes $B_{i1}$, $B_{i2}$,..., $B_{iH}$ de ses intervalles d'observation et de son point de censure $C_i$, soit que l'on a :

$$I\!P\left[T_i = t | X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] = I\!P\left[T_i = t | X^{i\infty}\right], \quad \forall\, t = 0, 1, 2, ... \tag{13}$$

autrement dit que, sachant $X^{i\infty}$, la connaissance des bornes $B_{i1}, ..., B_{iH}$ des intervalles d'observation et du point de censure $C_i$, c'est-à-dire que de la façon dont on observe en pratique le parcours de l'individu $i$, n'apporte aucune information sur sa durée effective (en pratique observée ou non de façon exacte) de chômage $T_i$. Cette hypothèse est automatiquement satisfaite si les bornes $B_{i1}, ..., B_{iH}$ et le point de censure $C_i$ sont fixes, non aléatoires, ou encore si ces valeurs sont des fonctions de $X^{i\infty}$.

La condition (13) d'indépendance conditionnelle entre $T_i$ et $(B_{i1}, ..., B_{iH}, C_i)$ implique (cf. preuve dans l'Annexe F) que :

$$\begin{aligned} & I\!P\left[t_1 \leq T_i \leq t_2 | X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] \\ = \; & I\!P\left[t_1 \leq T_i \leq t_2 | X^{i\infty}\right], \quad \forall\, t_1, t_2 = 0, 1, 2... \text{ (avec } t_1 \leq t_2) \end{aligned} \tag{14}$$

et

$$I\!P\left[T_i \geq t | X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] = I\!P\left[T_i \geq t | X^{i\infty}\right], \quad \forall\, t = 0, 1, 2, ... \tag{15}$$

et par conséquent aussi que[16] :

$$I\!P\left[T_i = t | T_i \geq t, X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] = I\!P\left[T_i = t | T_i \geq t, X^{i\infty}\right], \quad \forall\, t = 0, 1, 2, ...$$

Etant donné la condition (13) d'indépendance conditionnelle entre $T_i$ et $(B_{i1}, ..., B_{iH}, C_i)$, et ses implications (14) et (15), on peut déduire la distribution conditionnelle à $(X^{i\infty}, B_{i1}, ..., B_{iH}, C_i)$ de la durée de chômage effectivement observée $T_i^*$ d'un individu $i$ pris au hasard. Elle est donnée par la fonction de

[16] En effet, étant donné (13) et (15) et le fait que $(T_i = t) \Rightarrow (T_i \geq t)$, $\forall\, t = 0, 1, 2, ...$, on a :

$$\begin{aligned} I\!P\left[T_i = t | T_i \geq t, X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] &= \frac{I\!P\left[T_i = t | X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right]}{I\!P\left[T_i \geq t | X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right]} = \frac{I\!P\left[T_i = t | X^{i\infty}\right]}{I\!P\left[T_i \geq t | X^{i\infty}\right]} \\ &= I\!P\left[T_i = t | T_i \geq t, X^{i\infty}\right] \end{aligned}$$

densité conditionnelle :

$$
\begin{aligned}
& \overline{f}^{*}(t|X^{i\infty}, B_{i1}, ..., B_{iH}, C_i) \\
= \; & I\!P\left[T_i^* = t|X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right], \; \forall\, t = B_{i1}, ..., B_{iH}, B_{iH}+1, ..., C_i, C_i+1 \\
= \; & \begin{cases} \begin{array}{l} I\!P[\underline{t} \leq T_i \leq t|X^{i\infty}, B_{i1}, ..., B_{iH}, C_i] \\ = I\!P[\underline{t} \leq T_i \leq t|X^{i\infty}] \end{array} & , \begin{array}{l} \text{si } t = B_{ih}, \text{ où } \underline{t} = B_{ih-1}+1 \\ (\forall\, h = 1, 2, ..., H) \end{array} \\ I\!P[T_i = t|X^{i\infty}, B_{i1}, ..., B_{iH}, C_i] = I\!P[T_i = t|X^{i\infty}] & , \text{si } B_{iH}+1 \leq t \leq C_i \\ I\!P[T_i \geq t|X^{i\infty}, B_{i1}, ..., B_{iH}, C_i] = I\!P[T_i \geq t|X^{i\infty}] & , \text{si } t = C_i + 1 \end{cases}
\end{aligned}
$$

qui, puisqu'étant donné l'hypothèse d'exogénéité stricte des $X_{it}$, si le modèle est bien correctement spécifié, on a :

$$
\begin{aligned}
I\!P\left[T_i = t|X^{i\infty}\right] &= I\!P\left[T_i = t|X^{it}\right] = f(t, X^{it}; \beta^o) \\
&= \lambda(t, X^{it}; \beta^o) S(t, X^{it}; \beta^o), \;\; \forall\, t = 0, 1, 2, ...
\end{aligned}
$$

et[17]

$$
\begin{aligned}
& I\!P\left[\underline{t} \leq T_i \leq t|X^{i\infty}\right] = I\!P\left[\underline{t} \leq T_i \leq t|X^{it}\right] \\
= \; & \sum_{k^*=\underline{t}}^{t} I\!P\left[T_i = k^*|X^{ik^*}\right] = \sum_{k^*=\underline{t}}^{t} f(k^*, X^{ik^*}; \beta^o) \\
= \; & \sum_{k^*=\underline{t}}^{t} \lambda(k^*, X^{ik^*}; \beta^o) S(k^*, X^{ik^*}; \beta^o), \;\; \forall\, \underline{t}, t = 0, 1, 2... \text{ (avec } \underline{t} \leq t)
\end{aligned}
$$

et encore

$$
\begin{aligned}
I\!P\left[T_i \geq t|X^{i\infty}\right] &= I\!P\left[T_i \geq t|X^{it}\right] \\
&= S(t, X^{it}; \beta^o), \;\; \forall\, t = 0, 1, 2, ...
\end{aligned}
$$

et que, étant donné (6)[18], on a également :

$$
\begin{aligned}
S(t, X^{it}; \beta^o) &= I\!P\left[T_i \geq t|X^{it}\right] = I\!P\left[T_i \geq t|X^{it-1}\right] \\
&= S(t, X^{it-1}; \beta^o) = \prod_{t^*=0}^{t-1} \left(1 - \lambda(t^*, X^{it^*}; \beta^o)\right), \;\; \forall\, t = 1, 2, ... \quad (16)
\end{aligned}
$$

[17] Car, étant donné l'hypothèse d'exogénéité de $X_{it}$, on a :

$$
I\!P\left[\underline{t} \leq T_i \leq t|X^{i\infty}\right] = \sum_{k^*=\underline{t}}^{t} I\!P\left[T_i = k^*|X^{i\infty}\right] = \sum_{k^*=\underline{t}}^{t} I\!P\left[T_i = k^*|X^{ik^*}\right]
$$

[18] C'est-à-dire que la fonction de survie $S(t, X^{it})$ ne dépend en fait que de $X^{it-1}$.

est égale à :

$$
\begin{aligned}
& \overline{f}^*(t|X^{i\infty}, B_{i1}, ..., B_{iH}, C_i) = I\!P\left[T_i^* = t|X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] \\
= \ & I\!P\left[T_i^* = t|X^{im_i^*}, B_{i1}, ..., B_{iH}, C_i\right], \ \ \forall\, t = B_{i1}, ..., B_{iH}, B_{iH}+1, ..., C_i, C_i+1 \\
= \ & f^*(t, X^{im_i^*}, B_{i1}, ..., B_{iH}, C_i; \beta^o) \\
= \ & \begin{cases} \sum_{k^*=\underline{t}}^{t} \lambda(k^*, X^{ik^*}; \beta^o) S(k^*, X^{ik^*}; \beta^o) & , \begin{array}{l} \text{si } t = B_{ih}, \text{ où } \underline{t} = B_{ih-1}+1 \\ (\forall\ h = 1, 2, ..., H) \end{array} \\ \lambda(t, X^{it}; \beta^o) S(t, X^{it}; \beta^o) & , \text{si } B_{iH}+1 \leq t \leq C_i \\ S(t, X^{it-1}; \beta^o) & , \text{si } t = C_i+1 \end{cases}
\end{aligned}
$$

où $m_i^* = \min(t, C_i)$ et pour rappel par convention $B_{i0} = -1$.

Pour un échantillon de $n$ individus tirés au hasard[19], les $n$ $(H+3)$-uples $(T_1^*, X^{1\infty}, B_{11}, ..., B_{1H}, C_1), ...., (T_n^*, X^{n\infty}, B_{n1}, ..., B_{nH}, C_n)$ sont indépendants et de même distribution, et un estimateur MVc $\hat{\beta}_n$ de $\beta^o$ est dès lors donné par :

$$
\begin{aligned}
\hat{\beta}_n \ = \ & \text{Argmax}_{\beta\in\Theta} \sum_{i=1}^{n} \ln f^*(T_i^*, X^{iM_i^*}, B_{i1}, ..., B_{iH}, C_i; \beta) \\
= \ & \text{Argmax}_{\beta\in\Theta} \sum_{i=1}^{n} \Bigg[ \mathbb{I}_{(T_i^*=B_{i1})} \ln\left( \sum_{k^*=0}^{T_i^*} \lambda(k^*, X^{ik^*}; \beta^o) S(k^*, X^{ik^*}; \beta) \right) \\
& + \mathbb{I}_{(T_i^*=B_{i2})} \ln\left( \sum_{k^*=B_{i1}+1}^{T_i^*} \lambda(k^*, X^{ik^*}; \beta^o) S(k^*, X^{ik^*}; \beta) \right) \\
& \vdots \\
& + \mathbb{I}_{(T_i^*=B_{iH})} \ln\left( \sum_{k^*=B_{iH-1}+1}^{T_i^*} \lambda(k^*, X^{ik^*}; \beta^o) S(k^*, X^{ik^*}; \beta) \right) \\
& + \mathbb{I}_{(B_{iH}+1\leq T_i^*\leq C_i)} \left( \ln \lambda(T_i^*, X^{iT_i^*}, \beta) + \ln S(T_i^*, X^{iT_i^*}, \beta) \right) \\
& + \mathbb{I}_{(T_i^*=C_i+1)} \ln S(T_i^*, X^{iT_i^*-1}, \beta) \Bigg] \\
= \ & \text{Argmax}_{\beta\in\Theta} \sum_{i=1}^{n} \left[ \sum_{h=1}^{H} \mathbb{I}_{(T_i^*=B_{ih})} \ln\left( \sum_{k^*=B_{ih-1}+1}^{T_i^*} \lambda(k^*, X^{ik^*}; \beta^o) S(k^*, X^{ik^*}; \beta) \right) \right] \\
& + \sum_{i=1}^{n} \mathbb{I}_{(B_{iH}+1\leq T_i^*\leq C_i)} \left( \ln \lambda(T_i^*, X^{iT_i^*}, \beta) + \ln S(T_i^*, X^{iT_i^*}, \beta) \right) \\
& + \sum_{i=1}^{n} \mathbb{I}_{(T_i^*=C_i+1)} \ln S(T_i^*, X^{iT_i^*-1}, \beta) \qquad (17)
\end{aligned}
$$

[19] Au sens strict, avec remise.

qu'on peut encore écrire :

$$
\begin{aligned}
\hat{\beta}_n &= \text{Argmax}_{\beta\in\Theta} \sum_{i=1}^{n} \ln f^*(T_i^*, X^{iM_i^*}, B_{i1}, ..., B_{iH}, C_i; \beta) \\
&= \text{Argmax}_{\beta\in\Theta} \sum_{i=1}^{n} \left[ \sum_{h=1}^{H} \mathbb{I}_{(T_i^*=B_{ih})} \ln \left( \sum_{k^*=B_{ih-1}+1}^{T_i^*} \lambda(k^*, X^{ik^*}; \beta^o) S(k^*, X^{ik^*}; \beta^o) \right) \right] \\
&\quad + \sum_{i=1}^{n} \mathbb{I}_{(B_{iH}+1\leq T_i^*\leq C_i)} \ln \lambda(T_i^*, X^{iT_i^*}, \beta) \\
&\quad + \sum_{i=1}^{n} \mathbb{I}_{(T_i^*\geq B_{iH}+1)} \ln S(T_i^*, X^{iM_i^*}, \beta)
\end{aligned}
$$

où $M_i^* = \min(T_i^*, C_i)$, les $\mathbb{I}_{(.)}$ sont des variables indicatrices, pour rappel[20] $B_{i0} = -1$, et :

$$
S(k^*, X^{ik^*}, \beta) = \begin{cases} 1 & , \text{si } k^* = 0 \\ \prod_{t^*=0}^{k^*-1} \left(1 - \lambda(k^*, X^{ik^*}; \beta)\right) & , \forall\, k^* = 1, 2, ... \end{cases}
$$

et

$$
\ln S(T_i^*, X^{iM_i^*}, \beta) = \begin{cases} 0 & , \text{si } T_i^* = 0 \\ \sum_{t^*=0}^{T_i^*-1} \ln\left(1 - \lambda(t^*, X^{it^*}; \beta)\right) & , \forall\, T_i^* = 1, 2, ..., C_i + 1 \end{cases}
$$

- De façon semblable aux cas des deux schémas d'échantillonnage examinés précédemment, on constate que l'estimateur MVc $\hat{\beta}_n$ de $\beta^o$ ne dépend pas de l'ensemble des variables de l'échantillon $((T_1^*, X^{1\infty}, B_{11}, ..., B_{1H}, C_1), ....,$ $(T_n^*, X^{n\infty}, B_{n1}, ..., B_{nH}, C_n))$, mais, étant donné l'hypothèse d'exogénéité stricte des $X_{it}$ (et (16)), seulement de $((T_1^*, X^{1M_1^*}, B_{11}, ..., B_{1H}, C_1), ...., (T_n^*, X^{nM_n^*}, B_{n1}$ $, ..., B_{nH}, C_n))$. Autrement dit, pour estimer $\beta^o$ par l'estimateur MVc $\hat{\beta}_n$, il n'est pas nécessaire de disposer d'observations de $((T_1^*, X^{1\infty}, B_{11}, ..., B_{1H}, C_1), ....,$ $(T_n^*, X^{n\infty}, B_{n1}, ..., B_{nH}, C_n))$, mais seulement d'observations de $((T_1^*, X^{1M_1^*}, B_{11},$ $..., B_{1H}, C_1), ...., (T_n^*, X^{nM_n^*}, B_{n1}, ..., B_{nH}, C_n))$. $\hat{\beta}_n$ est donc un estimateur MVc adéquat pour estimer $\beta^o$ sur base des observations dont on dispose dans le schéma d'échantillonnage supposé.
- L'estimateur $\hat{\beta}_n$ étant un estimateur MVc standard, sous des conditions de régularité générales, il possède les propriétés habituelles des estimateurs MVc, à savoir (si, comme supposé, le modèle est bien correctement spécifié et les $X_{it}$

[20] Rappelons également que le présent schéma d'échantillonnage suppose que l'on a toujours $B_{i1} < B_{i2} <$ $... < B_{ih} < ... < B_{iH-1} < B_{iH} < C_i$.

strictement exogènes) :

$$\hat{\beta}_n \xrightarrow{p} \beta^o \quad \text{et} \quad V^{o\frac{1}{2}}\sqrt{n}(\hat{\beta}_n - \beta^o) \xrightarrow{d} \mathcal{N}(0, I_p)$$

où[21] :

$$\begin{aligned} V^o &= \frac{1}{n}\sum_{i=1}^{n} E\left[\left(\frac{\partial \ln f^*(T_i^*, X^{iM_i^*}, B_{i1}, ..., B_{iH}, C_i; \beta^o)}{\partial \beta}\right)\left(\frac{\partial f^*(T_i^*, X^{iM_i^*}, B_{i1}, ..., B_{iH}, C_i; \beta^o)}{\partial \beta}\right)'\right] \\ &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial^2 \ln f^*(T_i^*, X^{iM_i^*}, B_{i1}, ..., B_{iH}, C_i; \beta^o)}{\partial \beta \partial \beta'}\right] \end{aligned}$$

soit en termes d'approximation utilisable est échantillon fini :

$$\hat{\beta}_n \approx \mathcal{N}(\beta^o, \hat{V}_n^{-1}/n)$$

avec

$$\hat{V}_n = -\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2 \ln f^*(T_i^*, X^{iM_i^*}, B_{i1}, ..., B_{iH}, C_i; \hat{\beta}_n)}{\partial \beta \partial \beta'}$$

- Remarque :

Il est facile, en suivant le même raisonnement que celui développé ci-dessus, de dériver l'estimateur MVc $\hat{\beta}_n$ de $\beta^o$ pour d'autres schémas d'échantillonnage avec censure à droite et observation (partielle) par intervalles que celui examiné ici : par exemple des variantes du présent schéma où, avant le point de censure, la durée $T_i$ n'est toujours observée que par intervalles (cas où on a $B_{iH} = C_i$, $\forall\, i$), ou encore où la durée $T_i$ est d'abord observée de façon complète, puis par intervalles, puis encore de façon complète jusqu'au point de censure, etc.... Dans tous les cas, on obtiendra (en supposant toujours qu'il y a indépendance conditionnelle entre, d'une part $T_i$, et d'autre part les bornes $B_{ih}$ des intervalles d'observations et le point de censure $C_i$, un estimateur MVc $\hat{\beta}_n$ de $\beta^o$ dont la fonction de log-vraisemblance $\sum_{i=1}^{n} \ln f^*(.)$ contient des termes du type du premier terme de (17) pour les durées observées par intervalles, du type du deuxième terme de (17) pour les durées observées de façon exacte, et du type du troisième et dernier terme de (17) pour les observations de durée censurées.

[21] On notera qu'à nouceau la matrice de variance-covariance $V^o$ ne dépend pas de $n$. Cela résulte simplement du fait que, comme l'échantillon $((T_1^*, X^{1\infty}, B_{11}, ..., B_{1H}, C_1), ..., (T_n^*, X^{n\infty}, B_{n1}, ..., B_{nH}, C_n))$ est i.i.d., on a :

$$\begin{aligned} & E\left[\left(\frac{\partial \ln f^*(T_i^*, X^{iM_i^*}, B_{i1}, ..., B_{iH}, C_i; \beta^o)}{\partial \beta}\right)\left(\frac{\partial f^*(T_i^*, X^{iM_i^*}, B_{i1}, ..., B_{iH}, C_i; \beta^o)}{\partial \beta}\right)'\right] \\ =\ & -E\left[\frac{\partial^2 \ln f^*(T_i^*, X^{iM_i^*}, B_{i1}, ..., B_{iH}, C_i; \beta^o)}{\partial \beta \partial \beta'}\right] = V^o, \quad \forall\, i = 1, ..., n \end{aligned}$$

#### 1.3.2. Estimation sans l'hypothèse d'exogénéité stricte des $X_{it}$

**A. Echantillonnage aléatoire simple**

- En principe, l'estimateur standard (pour $X_{it}$ exogène) reste adéquat : il est toujours convergent et asymptotiquement normal (avec la même matrice de variance-covariance asymptotique), mais correspond maintenant à un estimateur MVc partiel (et non MVc standard).

**B. Echantillonnage aléatoire avec censure à droite**

- En principe, l'estimateur standard (pour $X_{it}$ exogène) reste à nouveau adéquat : il est toujours convergent et asymptotiquement normal (avec la même matrice de variance-covariance asymptotique), mais correspond maintenant à un estimateur MVc partiel (et non MVc standard).

**C. Echantillonnage aléatoire avec censure à droite et observation (partielle) par intervalles**

- A priori, l'estimateur standard (pour $X_{it}$ exogène) n'est plus adéquat.

# 2. Modèle avec plusieurs destinations

On considère à présent une généralisation du modèle de base où les individus peuvent passer d'un état initial à plusieurs autres états (destinations multiples) mutuellement exclusifs : de façon concrète, par exemple le passage de l'état "chômeur" aux états "a trouvé un travail à temps plein", "a trouvé un travail à temps partiel", "a quitté la population active", etc... Cette généralisation revient simplement à décomposer la sortie de l'état initial, i.e. l'état "est sorti du chômage" du modèle de base, en plusieurs "sous-états" mutuellement exclusifs. On suppose que ces "sous-états" mutuellement exclusifs forment une partition, autrement dit que les diverses destinations considérées recouvrent toutes les destinations possibles.

## 2.1. Variables et notations

- A côté de la variable $T_i$ = durée de chômage, mesurée en mois, semaines ou jours, d'un individu $i$ pris au hasard dans une population donnée, dont les valeurs possibles sont supposées être $t = 0, 1, 2, ..., \infty$, on a maintenant une variable supplémentaire $E_i$ = destination vers laquelle l'individu $i$ se dirige (emploi temps plein, temps partiel, sortie population active, etc..) lorsqu'il sort, après une durée $T_i$, du chômage. On suppose que les valeurs possibles de $E_i$ sont $l = 1, 2, ..., L$, où $L$ est le nombre de destinations mutuellement exclusives vers lesquelles les individus peuvent sortir. On retrouve le modèle de base lorsque $L = 1$.
- Les conventions de notations (i.e. $X_{it}$, $X^{it}$, $X^{\underline{it+1}}$, $X^{i\infty}$) concernant les variables explicatives caractérisant l'individu $i$ et les conditions du marché du travail

rencontrées par l'individu $i$ au cours des différentes périodes $t$ restent inchangées.

- On cherche maintenant à modéliser le couple de variables $(T_i, E_i)$ durée de chômage/destination de sortie d'un individu $i$ pris au hasard au fonction des variables explicatives $X_{it}$ qui le caractérisent.

## 2.2. Ingrédients du modèle

### 2.2.1. Fonctions de hasard spécifique conditionnelles et fonction de hasard conditionnelle

- Les fonctions de hasard spécifique conditionnelles[22] sont définies par :

$$\lambda_l(t, X^{it}) = I\!P\left[T_i = t, E_i = l | T_i \geq t, X^{it}\right], \quad \begin{array}{l} \forall\, t = 0, 1, 2, ... \\ \forall\, l = 1, 2, ..., L \end{array} \tag{18}$$

Elle donnent la probabilité jointe qu'un individu $i$ pris au hasard ait une durée $T_i$ de chômage (exactement) égale à $t$ et que sa destination $E_i$ de sortie du chômage soit égale à $l$, sachant que sa durée $T_i$ de chômage est au moins égale à $t$ et les variables explicatives $X^{it}$. En termes de population, la probabilité $\lambda_l(t, X^{it})$ s'interprète comme la proportion des individus qui sortent du chômage en $t$ vers la destination $l$, parmi les individus de la population qui sont toujours au chômage en $t-1$ et qui ont les caractéristiques décrites par $X^{it}$.

- Les fonctions de hasard spécifique conditionnelles $\lambda_l(t, X^{it})$, $l = 1, 2, ..., L$ forment le coeur du modèle. Ce sont les fonctions que l'on cherche à estimer. On notera que ces fonctions de hasard spécifique conditionnelles $\lambda_l(t, X^{it})$ ne constituent pas des fonctions de hasard au sens habituel du terme.

- Comme dans le modèle de base, la fonction de hasard conditionnelle est définie par :

$$\lambda(t, X^{it}) = I\!P\left[T_i = t | T_i \geq t, X^{it}\right], \;\; \forall\, t = 0, 1, 2, ...$$

C'est la probabilité que, quelque soit sa destination $E_i$ de sortie, la durée $T_i$ de chômage d'un individu $i$ pris au hasard soit (exactement) égale à $t$, sachant qu'elle est au moins égale à $t$ et les variables $X^{it}$. En termes de population, cette probabilité $\lambda(t, X^{it})$ s'interprète comme la proportion des individus qui sortent du chômage en (exactement) $t$, quelle que soit leur destination $l$, parmi les individus de la population qui sont toujours au chômage en $t-1$ et qui ont les caractéristiques décrites par $X^{it}$.

- Les différentes sorties $l = 1, 2, ..., L$ formant une partition, la fonction de hasard conditionnelle $\lambda(t, X^{it})$ est tout simplement égale à la somme des fonctions de

[22] ainsi appelées par analogie aux modèles (latents, en temps continu) à risques concurrents (competing risk models).

hasard spécifique conditionnelles $\lambda_l(t, X^{it})$ pour les différentes sorties[23] :

$$\lambda(t, X^{it}) = \sum_{l=1}^{L} \lambda_l(t, X^{it}), \ \ \forall\, t = 0, 1, 2, ... \tag{19}$$

### 2.2.2. Exogénéité stricte des $X_{it}$

- Dans ce modèle plus général, on dit que les variables explicatives $X_{it}$ sont strictement exogènes si[24] :

$$D\left(X^{\underline{it+1}}|T_i = t-k, E_i = l, X^{it}\right) = D\left(X^{\underline{it+1}}|X^{it}\right), \quad \begin{array}{l} \forall\, t = 0, 1, 2, ... \\ \forall\, k = 0, 1, 2, ..., t \\ \forall\, l = 1, 2, ..., L \end{array} \tag{20}$$

  où $D\,(.|.)$ désigne les distributions conditionnelles de $X^{\underline{it+1}}$ sachant respectivement $(T_i = t-k, E_i = l, X^{it})$ et $X^{it}$, autrement dit si, sachant $X^{it}$, la connaissance que la durée $T_i$ est égale à $t-k$ ($\forall\, k = 0, 1, 2, ..., t$) et que la destination de sortie $E_i$ est égale à $l$ ($\forall\, l = 1, 2, ..., L$), n'apporte aucune information nouvelle pour prédire $X^{\underline{it+1}}$.

- Remarques :

  a- Comme dans le cas du modèle de base, la condition d'exogénéité stricte des $X_{it}$ est trivialement satisfaite lorsque les variables explicatives $X_{it}$ sont invariantes dans le temps. On peut en pratique souvent, mais pas toujours, considérer cette condition d'exogénéité comme satisfaite.

  b- La condition (20) d'exogénéité stricte des $X_{it}$ implique la condition (2) d'exogénéité stricte des $X_{it}$ utilisée dans le cadre du modèle de base, c-à-d. :

$$D\left(X^{\underline{it+1}}|T_i = t-k, X^{it}\right) = D\left(X^{\underline{it+1}}|X^{it}\right), \quad \begin{array}{l} \forall\, t = 0, 1, 2, ... \\ \forall\, k = 0, 1, 2, ..., t \end{array} \tag{21}$$

  et s'y réduit dans le cas où $L = 1$.

- La condition (20) d'exogénéité stricte des $X_{it}$ implique (cf. preuve dans l'Annexe

---

[23] Formellement :

$$\begin{aligned} \lambda(t, X^{it}) &= I\!P\left[T_i = t | T_i \geq t, X^{it}\right] = I\!P\left[T_i = t, (E_i = 1 \text{ ou ... ou } E_i = L)\, | T_i \geq t, X^{it}\right] \\ &= I\!P\left[(T_i = t, E_i = 1) \text{ ou ... ou } (T_i = t, E_i = L) | T_i \geq t, X^{it}\right] \\ &= \sum_{l=1}^{L} I\!P\left[T_i = t, E_i = l | T_i \geq t, X^{it}\right] = \sum_{l=1}^{L} \lambda_l(t, X^{it}) \end{aligned}$$

[24] On notera que, de façon semblable au cas du modèle de base, cette condition d'exogénéité peut être exprimée de façon équivalente et plus simple, mais moins intuitive par (cf. preuve dans l'annexe G):

$$D\left(X^{\underline{it+1}}|T_i = t, E_i = l, X^{it}\right) = D\left(X^{\underline{it+1}}|X^{it}\right), \ \ \forall\, t = 0, 1, 2, ..., \ \forall\, l = 1, 2, ..., L$$

H) que :

$$\begin{aligned} \mathbb{P}\left[T_i = t, E_i = l | X^{it}, X^{\underline{it+1}}\right] &= \mathbb{P}\left[T_i = t, E_i = l | X^{i\infty}\right] \\ &= \mathbb{P}\left[T_i = t, E_i = l | X^{it}\right], \quad \begin{array}{l} \forall\, t = 0,1,2,... \\ \forall\, l = 1,2,...,L \end{array} \end{aligned} \tag{22}$$

autrement dit que, sachant $X^{it}$, la connaissance de $X^{\underline{it+1}}$ n'apporte aucune information sur la probabilité jointe qu'un individu $i$ pris au hasard ait une durée $T_i$ de chômage égale à $t$ et que sa destination $E_i$ de sortie du chômage soit égale à $l$.

Comme dans le modèle de base (puisque (20) $\Rightarrow$ (21)), elle implique aussi que :

$$\begin{aligned} \mathbb{P}\left[T_i = t | X^{it}, X^{\underline{it+1}}\right] &= \mathbb{P}\left[T_i = t | X^{i\infty}\right] \\ &= \mathbb{P}\left[T_i = t | X^{it}\right], \quad \forall\, t = 0,1,2,... \end{aligned} \tag{23}$$

$$\begin{aligned} \mathbb{P}\left[T_i \geq t | X^{it}, X^{\underline{it+1}}\right] &= \mathbb{P}\left[T_i \geq t | X^{i\infty}\right] \\ &= \mathbb{P}\left[T_i \geq t | X^{it}\right], \quad \forall\, t = 0,1,2,... \end{aligned} \tag{24}$$

et

$$\begin{aligned} \mathbb{P}\left[T_i = t | T_i \geq t, X^{it}, X^{\underline{it+1}}\right] &= \mathbb{P}\left[T_i = t | T_i \geq t, X^{i\infty}\right] \\ &= \mathbb{P}\left[T_i = t | T_i \geq t, X^{it}\right] \\ &= \lambda(t, X^{it}), \quad \forall\, t = 0,1,2,... \end{aligned} \tag{25}$$

autrement dit que, sachant $X^{it}$, la connaissance de $X^{\underline{it+1}}$ n'apporte aucune information sur la probabilité que, quelle que soit sa destination de sortie $E_i$, la durée $T_i$ de chômage d'un individu $i$ pris au hasard soit égale à $t$, supérieure ou égale à $t$, ou encore égale à $t$ sachant qu'elle est supérieure ou égale à $t$.

Enfin, elle implique encore que :

$$\begin{aligned} \mathbb{P}\left[T_i = t, E_i = l | T_i \geq t, X^{it}, X^{\underline{it+1}}\right] &= \mathbb{P}\left[T_i = t, E_i = l | T_i \geq t, X^{i\infty}\right] \\ &= \mathbb{P}\left[T_i = t, E_i = l | T_i \geq t, X^{it}\right] \\ &= \lambda_l(t, X^{it}), \quad \begin{array}{l} \forall\, t = 0,1,2,... \\ \forall\, l = 1,2,...,L \end{array} \end{aligned} \tag{26}$$

autrement dit que, sachant $X^{it}$, la connaissance de $X^{\underline{it+1}}$ n'apporte aucune information sur la probabilité jointe que la destination $E_i$ de sortie d'un individu $i$ pris au hasard soit égale à $l$ et que sa durée $T_i$ de chômage soit égale à $t$, sachant que cette durée est supérieure ou égale à $t$.

### 2.2.3. Fonction de survie conditionnelle

- Comme dans le modèle de base, le fonction de survie conditionnelle est définie par :

$$S(t, X^{it}) = \mathbb{P}\left[T_i \geq t | X^{it}\right], \quad \forall\, t = 0,1,2,...$$

C'est la probabilité que, quelque soit sa destination $E_i$ de sortie, la durée $T_i$

de chômage d'un individu $i$ pris au hasard soit supérieur ou égale à $t$, sachant les variables $X^{it}$. En termes de population, cette probabilité $S(t, X^{it})$ s'interprète comme la proportion des individus qui, quelle que soit leur destination de $E_i$ de sortie, sortent du chômage en $t$ ou après $t$, ou ce qui revient au même la proportion des individus qui sont toujours au chômage en $t-1$, parmi les individus de la population qui ont les caractéristiques décrites par $X^{it}$.

- Comme dans le modèle de base, sous l'hypothèse d'exogénéité stricte des $X_{it}$, étant donné (24), on a :

  $$S(t, X^{it}) = I\!P\left[T_i \geq t | X^{it}\right] = I\!P\left[T_i \geq t | X^{i\infty}\right] = \overline{S}(t|X^{i\infty}), \;\; \forall\, t = 0, 1, 2, ...$$

  où la fonction $\overline{S}(t|X^{i\infty})$ est la fonction de survie (conditionnelle) associée à la distribution conditionnelle définie par :

  $$\overline{f}(t|X^{i\infty}) = I\!P\left[T_i = t | X^{i\infty}\right], \;\; \forall\, t = 0, 1, 2, ...$$

  c'est-à-dire la distribution conditionnelle de $T_i$ sachant $X^{i\infty}$. Autrement dit, sous l'hypothèse d'exogénéité stricte des $X_{it}$, la fonction $S(t, X^{it})$ défini bien, comme son nom le suggère, une fonction de survie (conditionnelle) au sens habituel du terme, découlant d'une distribution conditionnelle bien définie. Par contre, à nouveau comme dans le modèle de base, si on renonce à l'hypothèse d'exogénéité stricte des $X_{it}$, la fonction $S(t, X^{it})$ ne peut plus être interprétée comme une fonction de survie (conditionnelle) au sens habituel du terme, découlant d'une distribution conditionnelle bien définie, mais seulement comme une collection (pour $t = 0, 1, 2,...$) de probabilités conditionnelles.

- Toujours comme dans le modèle de base, sous l'hypothèse d'exogénéité stricte des $X_{it}$, de la fonction de hasard conditionnelle $\lambda(t, X^{it})$, on peut déduire la fonction de survie conditionnelle $S(t, X^{it})$. La condition (20) d'exogénéité stricte des $X_{it}$ impliquant la condition (21) d'exogénéité stricte des $X_{it}$ utilisée dans le cadre du modèle de base, les résultats obtenu sous cette hypothèse dans le modèle de base restent en effet d'application dans le cadre du modèle à destinations multiples. Et comme la fonction de hasard conditionnelle $\lambda(t, X^{it})$ est simplement égale à la somme des fonctions de hasard spécifique conditionnelles $\lambda_l(t, X^{it})$ (cf. équation (19)), on peut en fait écrire la fonction de survie conditionnelle $S(t, X^{it})$ comme une fonction des fonctions de hasard spécifique conditionnelles $\lambda_l(t, X^{it})$. Ainsi, en substituant (19) dans l'expression dérivée à la Section 1.2.3 pour $S(t, X^{it})$ sous l'hypothèse d'exogénéité stricte des $X_{it}$, on obtient :

  $$S(t, X^{it}) = I\!P\left[T_i \geq t | X^{it}\right] = \begin{cases} 1 & , \text{si } t = 0 \\ \displaystyle\prod_{t^*=0}^{t-1} \left(1 - \sum_{l=1}^{L} \lambda_l(t^*, X^{it^*})\right) & , \forall\, t = 1, 2, ... \end{cases}$$

- Remarque :

  On notera à nouveau que dans la formule ci-dessus, $S(t, X^{it})$ ne dépend en fait que de $X^{it-1} = \left(X'_{i0}, X'_{i1}, ..., X'_{it-1}\right)'$. Autrement dit, sous l'hypothèse d'exogénéité stricte des $X_{it}$, on a en fait[25] :

$$\begin{aligned} S(t, X^{it}) &= I\!P\left[T_i \geq t | X^{it}\right] = I\!P\left[T_i \geq t | X^{it-1}\right] \\ &= S(t, X^{it-1}) = \prod_{t^*=0}^{t-1} \left(1 - \sum_{l=1}^{L} \lambda_l(t^*, X^{it^*})\right), \ \forall\, t = 1, 2, ... \end{aligned} \tag{27}$$

- Si on renonce l'hypothèse d'exogénéité stricte des $X_{it}$, on peut toujours exprimer la fonction $S(t, X^{it})$ par rapport aux fonctions de hasard spécifique conditionnelles $\lambda_l(t, X^{it})$, mais, comme pour le modèle de base, elle n'a plus la forme simple donnée ci-dessus. En substituant comme précédemment (19) dans l'expression dérivée à la Section 1.2.3 pour $S(t, X^{it})$ pour cette fois le cas où l'hypothèse d'exogénéité stricte des $X_{it}$ ne tient pas, on voit qu'elle prend la forme plus complexe :

$$S(t, X^{it}) = I\!P\left[T_i \geq t | X^{it}\right] = \begin{cases} 1 & , \text{si } t = 0 \\ \prod_{t^*=0}^{t-1} \left(1 - \sum_{l=1}^{L} \lambda_l(t^*, X^{it^*})\right) \\ \times \prod_{t^*=0}^{t-1} \dfrac{I\!P\left[X_{it^*+1} | T_i \geq t^* + 1, X^{it^*}\right]}{I\!P\left[X_{it^*+1} | X^{it^*}\right]} & , \forall\, t = 1, 2, ... \end{cases}$$

forme qui dépend notamment de la distribution des $X_{it}$.. Ainsi, si on renonce à l'hypothèse d'exogénéité stricte des $X_{it}$, non seulement la fonction $S(t, X^{it})$ ne définit pas une fonction de survie au sens habituel du terme, et ne peut donc pas être interprétée comme telle, mais son expression en termes de fonctions de hasard spécifique conditionnelles $\lambda_l(t, X^{it})$ n'est plus aussi simple que dans le cas où les $X_{it}$ sont strictement exogènes et dépend notamment de la distribution des $X_{it}$.

### 2.2.4. Fonction de densité jointe conditionnelle et fonction de densité conditionnelle

- La fonction de densité jointe conditionnelle est définie par :

$$f_L(t, l, X^{it}) = I\!P\left[T_i = t, E_i = l | X^{it}\right], \quad \begin{array}{l} \forall\, t = 0, 1, 2, ... \\ \forall\, l = 1, 2, ..., L \end{array}$$

C'est la probabilité jointe qu'un individu $i$ pris au hasard ait une durée $T_i$ de chômage (exactement) égale à $t$ et que sa destination $E_i$ de sortie du chômage soit égale à $l$, sachant les variables explicatives $X^{it}$. En termes de population, le probabilité $f_L(t, l, X^{it})$ s'interprète comme la proportion des individus qui sortent du chômage en $t$ vers la destination $l$, parmi les individus de la population qui ont les caractéristiques décrites par $X^{it}$.

[25] On remarquera que cette égalité n'a aucun sens en $t = 0$.

- Sous l'hypothèse d'exogénéité stricte des $X_{it}$, étant donné (22), on a :

$$\begin{aligned} f_L(t,l,X^{it}) &= I\!P\left[T_i = t, E_i = l | X^{it}\right] \\ &= I\!P\left[T_i = t, E_i = l | X^{i\infty}\right] = \bar{f}_L(t, l | X^{i\infty}), \quad \begin{array}{l} \forall\, t = 0, 1, 2, ... \\ \forall\, l = 1, 2, ..., L \end{array} \end{aligned}$$

où $\bar{f}_L(t, l|X^{i\infty})$ défini la distribution jointe conditionnelle de $(T_i, E_i)$ sachant $X^{i\infty}$. Autrement dit, sous l'hypothèse d'exogénéité stricte des $X_{it}$, la fonction $f_L(t, l, X^{it})$ défini bien, comme son nom le suggère, une distribution (discrète) jointe conditionnelle au sens habituel du terme. Par contre, si on renonce à l'hypothèse d'exogénéité stricte des $X_{it}$, contrairement à ce que son nom suggère, la fonction $f_L(t, l, X^{it})$ ne peut plus être interprétée comme une distribution (discrète) jointe conditionnelle au sens habituel du terme, mais seulement, dans le modèle de base, comme une collection (pour $t = 0, 1, 2,...$) de probabilités jointes conditionnelles.

- Comme pour la fonction de survie conditionnelle, des fonctions de hasard spécifique conditionnelles $\lambda_l(t, X^{it})$, on peut déduire la fonction de densité jointe conditionnelle $f_L(t, l, X^{it})$. Par définition, de façon générale, on a[26] :

$$\begin{aligned} I\!P\left[T_i = t, E_i = l | X^{it}\right] &= I\!P\left[T_i = t, E_i = l | T_i \geq t, X^{it}\right] I\!P\left[T_i \geq t | X^{it}\right] \\ &= \lambda_l(t, X^{it}) S(t, X^{it}), \quad \begin{array}{l} \forall\, t = 0, 1, 2, ... \\ \forall\, l = 1, 2, ..., L \end{array} \end{aligned} \tag{28}$$

de sorte que, en utilisant les expressions obtenues pour $S(t, X^{it})$ à la section précédente, sous l'hypothèse d'exogénéité stricte des $X_{it}$, $\forall\, l = 1, 2, ..., L$, on a :

$$\begin{aligned} f_L(t,l,X^{it}) &= I\!P\left[T_i = t, E_i = l | X^{it}\right] \\ &= \begin{cases} \lambda_l(0, X^{i0}) & , \text{si } t = 0 \\ \lambda_l(t, X^{it}) \displaystyle\prod_{t^*=0}^{t-1} \left(1 - \sum_{l=1}^{L} \lambda_l(t^*, X^{it^*})\right) & , \forall\, t = 1, 2, ... \end{cases} \end{aligned}$$

tandis que si on renonce l'hypothèse d'exogénéité stricte des $X_{it}$, $\forall\, l = 1, 2, ..., L$, on a la forme plus complexe :

$$\begin{aligned} f_L(t,l,X^{it}) &= I\!P\left[T_i = t, E_i = l | X^{it}\right] \\ &= \begin{cases} \lambda_l(0, X^{i0}) & , \text{si } t = 0 \\ \lambda_l(t, X^{it}) \displaystyle\prod_{t^*=0}^{t-1} \left(1 - \sum_{l=1}^{L} \lambda_l(t^*, X^{it^*})\right) & \\ \times \displaystyle\prod_{t^*=0}^{t-1} \frac{I\!P\left[X_{it^*+1} | T_i \geq t^* + 1, X^{it^*}\right]}{I\!P\left[X_{it^*+1} | X^{it^*}\right]} & , \forall\, t = 1, 2, ... \end{cases} \end{aligned}$$

[26] La validité de la relation :

$$I\!P\left[T_i = t, E_i = l | X^{it}\right] = I\!P\left[T_i = t, E_i = l | T_i \geq t, X^{it}\right] I\!P\left[T_i \geq t | X^{it}\right]$$

vient, de façon semblable au cas de la note 3, du fait que l'événement $(T_i = t)$ implique l'événement $(T_i \geq t)$.

forme qui, comme on peut le voir, dépend à nouveau notamment de la distribution des $X_{it}$. Ainsi, on a encore que, si on renonce à l'hypothèse d'exogénéité stricte des $X_{it}$, non seulement la fonction $f_L(t,l,X^{it})$ ne définit pas une fonction de densité jointe au sens habituel du terme, et ne peut donc pas être interprétée comme telle, mais encore que son expression en termes des fonctions de hasard spécifique conditionnelles $\lambda_l(t,X^{it})$ n'est plus aussi simple que dans le cas où les $X_{it}$ sont strictement exogènes et dépend notamment de la distribution des $X_{it}$.

- Comme dans le modèle de base, la fonction de densité conditionnelle est définie par :

$$f(t,X^{it}) = I\!P\left[T_i = t|X^{it}\right], \ \forall\, t = 0,1,2,...$$

C'est la probabilité que, quelque soit sa destination $E_i$ de sortie, la durée $T_i$ de chômage d'un individu $i$ pris au hasard soit (exactement) égale à $t$, sachant les variables $X^{it}$. En termes de population, cette probabilité $f(t,X^{it})$ s'interprète comme la proportion des individus qui, quelle que soit leur destination de $E_i$ de sortie, sortent du chômage en (exactement) $t$ parmi les individus de la population qui ont les caractéristiques décrites par $X^{it}$.

- A nouveau comme dans le modèle de base, sous l'hypothèse d'exogénéité stricte des $X_{it}$, étant donné (23), on a :

$$f(t,X^{it}) = I\!P\left[T_i = t|X^{it}\right] = I\!P\left[T_i = t|X^{i\infty}\right] = \overline{f}(t|X^{i\infty}), \ \forall\, t = 0,1,2,...$$

où $\overline{f}(t|X^{i\infty})$ défini la distribution conditionnelle de $T_i$ sachant $X^{i\infty}$. Autrement dit, sous l'hypothèse d'exogénéité stricte des $X_{it}$, la fonction $f(t,X^{it})$ défini bien une distribution (discrète) conditionnelle au sens habituel du terme. Par contre, si on renonce à l'hypothèse d'exogénéité stricte des $X_{it}$, la fonction $f(t,X^{it})$ ne peut à nouveau plus être interprétée comme une distribution (discrète) conditionnelle au sens habituel du terme, mais seulement comme une collection (pour $t = 0,1,2,...$) de probabilités conditionnelles.

- Par définition, de façon générale, on a :

$$\begin{aligned} f(t,X^{it}) &= I\!P\left[T_i = t|X^{it}\right] = \sum_{l=1}^{L} I\!P\left[T_i = t, E_i = l|X^{it}\right] \\ &= \sum_{l=1}^{L} f_L(t,l,X^{it}), \ \forall\, t = 0,1,2,... \end{aligned}$$

soit, étant donné (28),

$$f(t,X^{it}) = I\!P\left[T_i = t|X^{it}\right] = \sum_{l=1}^{L} \lambda_l(t,X^{it})S(t,X^{it}), \ \forall\, t = 0,1,2,...$$

Des expressions obtenues précédemment, on déduit que sous l'hypothèse d'exogénéité stricte des $X_{it}$, on a :

$$f(t, X^{it}) = I\!P\left[T_i = t | X^{it}\right] = \begin{cases} \sum_{l=1}^{L} \lambda_l(0, X^{i0}) & , \text{si } t = 0 \\ \sum_{l=1}^{L} \lambda_l(t, X^{it}) \left[\prod_{t^*=0}^{t-1} \left(1 - \sum_{l=1}^{L} \lambda_l(t^*, X^{it^*})\right)\right] & , \forall\, t = 1, 2, ... \end{cases}$$

tandis que si on renonce l'hypothèse d'exogénéité stricte des $X_{it}$, on a la forme plus complexe :

$$f(t, X^{it}) = I\!P\left[T_i = t | X^{it}\right] = \begin{cases} \sum_{l=1}^{L} \lambda_l(0, X^{i0}) & , \text{si } t = 0 \\ \sum_{l=1}^{L} \lambda_l(t, X^{it}) \left[\prod_{t^*=0}^{t-1} \left(1 - \sum_{l=1}^{L} \lambda_l(t^*, X^{it^*})\right) \right. & \\ \left. \times \prod_{t^*=0}^{t-1} \frac{I\!P\left[X_{it^*+1} | T_i \geq t^* + 1, X^{it^*}\right]}{I\!P[X_{it^*+1} | X^{it^*}]} \right] & , \forall\, t = 1, 2, ... \end{cases}$$

forme qui dépend à nouveau notamment de la distribution des $X_{it}$. Ainsi, on a toujours que, si on renonce à l'hypothèse d'exogénéité stricte des $X_{it}$, non seulement la fonction $f(t, X^{it})$ ne définit pas une fonction de densité au sens habituel du terme, et ne peut donc pas être interprétée comme telle, mais encore que son expression en termes des fonctions de hasard spécifique conditionnelles $\lambda_l(t, X^{it})$ n'est plus aussi simple que dans le cas où les $X_{it}$ sont strictement exogènes et dépend notamment de la distribution des $X_{it}$.

On notera que, comme pour la fonction de survie et pour les mêmes raisons, les expressions données ci-dessus pour la fonction de densité conditionnelle $f(t, X^{it})$ sont identiques à celles obtenues dans le cadre du modèle de base : on a simpement que la fonction de densité conditionnelle $\lambda(t, X^{it})$ est exprimée comme la somme des fonctions de hasard spécifique conditionnelles $\lambda_l(t, X^{it})$.

## 2.3. Estimation du modèle

- On cherche à estimer les fonctions de hasard spécifique conditionnelles $\lambda_l(t, X^{it})$ d'un individu $i$ tiré au hasard parmi les individus d'une population cible (par exemple, les individus qui s'inscrivent comme demandeur d'emploi au cours d'une période donnée).

- On commence par spécifier des formes fonctionnelles $\lambda_l(t, X^{it}; \beta_l)$, qui dépendent de vecteurs $p_l \times 1$ de paramètres $\beta_l \in \Theta_l$, pour les fonctions de hasard spécifique conditionnelles $\lambda_l(t, X^{it})$, formes fonctionnelles que l'on suppose correctement spécifiées, i.e. telle que :

$$\lambda_l(t, X^{it}; \beta_l^o) = \lambda_l(t, X^{it}), \text{ pour des valeurs } \beta_l^o \in \Theta_l, \forall\, l = 1, 2, ..., L$$

- Les valeurs des fonctions de hasard spécifique conditionnelles doivent toujours être comprise[27] entre 0 et 1. De façon à ce que cette restriction soit toujours satisfaite, on spécifiera par exemple :

$$\lambda_l(t, X^{it}; \beta_l) = g\left(X^{it*\prime}\beta_l\right)$$

où $g(x) = \dfrac{e^x}{1+e^x}$ (= fonction (cdf.) logistique) et $X^{it*}$ est un vecteur de variables composé de $t$, (d'un sous-ensemble)[28] de $X^{it}$, et de leur puissances (carré, cube, ...) et interactions (autrement dit, une forme polynomiale en $t$ et (le sous-ensemble choisi de) $X^{it}$). Bien que cela ne garantisse seulement que la positivité des valeurs des fonctions de hasard spécifique conditionnelles, une autre spécification a priori intéressante (pour l'interprétation des paramètres $\beta_l$ en terme de semi-élasticité) pour le fonction $g(x)$ est $g(x) = e^x$. Des restrictions croisées entre les vecteurs de paramètres $\beta_l$ peuvent éventuellement être envisagées.

- Comme dans le modèle de base, contrairement aux fonctions de survie et de densité (jointe) conditionnelles, les fonctions de hasard spécifique conditionnelles ont toujours un sens précis et interprétable d'un point de vue pratique, que l'hypothèse d'exogénéité stricte des $X_{it}$ soit ou non satisfaite. Que cette hypothèse soit ou non satisfaite, il est donc toujours légitime de chercher à estimer les fonctions de hasard spécifique conditionnelles $\lambda_l(t, X^{it})$ d'un individu $i$ pris au hasard. Dans le cas où l'hypothèse d'exogénéité stricte des $X_{it}$ n'est pas satisfaite, on a simplement que, des fonctions de hasard spécifique conditionnelles, on ne peut pas déduire des fonctions de survie et de densité (jointe) conditionnelles au sens habituel de ces termes. Dans la suite, on considère le problème de l'estimation des paramètres des fonctions de hasard spécifique conditionnelles $\lambda_l(t, X^{it}; \beta_l)$ supposée correctement spécifiée, tout d'abord sous l'hypothèse d'exogénéité stricte des $X_{it}$, et ensuite sans faire appel à cette hypothèse.

### 2.3.1. Estimation sous l'hypothèse d'exogénéité stricte des $X_{it}$

- Si on suppose que les variables explicatives $X_{it}$ sont strictement exogènes, pour les formes fonctionnelles $\lambda_l(t, X^{it}; \beta_l)$ des fonctions de hasard spécifique conditionnelles, étant donné ce qui précède, on a comme forme fonctionnelle pour les fonctions de survie et de densité jointe conditionnelle :

$$S(t, X^{it}; \beta) = \begin{cases} 1 & , \text{si } t = 0 \\ \prod\limits_{t^*=0}^{t-1}\left(1 - \sum\limits_{l=1}^{L}\lambda_l(t^*, X^{it^*}; \beta_l)\right) & , \forall\, t = 1, 2, ... \end{cases} \tag{29}$$

[27] Les fonctions de hasard spécifique conditionnelles donnent des probabilités conditionnelles, et une probabilité conditionnelle est toujours comprise entre 0 et 1.

[28] En pratique, le vecteur de variables explicatives $X_{it}$ correspondant à la période $t$ considérée, dans lequel peuvent éventuellement être incorporées des variables explicatives retardées (avec un ou plusieurs retards).

et

$$f_L(t,l,X^{it};\beta) = \lambda_l(t,X^{it};\beta_l)S(t,X^{it};\beta), \quad \begin{array}{l}\forall\, t = 0,1,2,... \\ \forall\, l = 1,2,...,L\end{array}$$

où $\beta = (\beta_1', \beta_2', ..., \beta_L')'$ est un vecteur $p \times 1$ de paramètres ($p = \sum_{l=1}^{L} p_l$).

- Sous l'hypothèse que les formes fonctionnelles $\lambda_l(t, X^{it}; \beta_l$ sont correctement spécifiées, on a évidemment que $S(t, X^{it}; \beta)$ et $f_L(t, l, X^{it}; \beta)$ sont correctement spécifiées, i.e. que :

$$\left\{ \begin{array}{l} S(t,X^{it};\beta^o) = S(t,X^{it}) \\ f_L(t,l,X^{it};\beta^o) = f_L(t,l,X^{it}) \end{array} \right. , \quad \text{pour } \beta^o \in \Theta$$

où $\beta^o = (\beta_1^{o\prime}, \beta_2^{o\prime}, ..., \beta_L^{o\prime})'$ et $\Theta \equiv \Theta_1 \times \Theta_2 \times ... \times \Theta_L$ (si les vecteurs de paramètres $\beta_l$ sont non liés).

- Ayant spécifié les fonctions de hasard spécifique conditionnelles $\lambda_l(t, X^{it}; \beta_l)$ dont découlent, sous l'hypothèse d'exogénéité stricte des $X_{it}$, des spécifications pour les fonctions de survie et de densité jointe conditionnelles, on peut, comme dans le modèle de base, en supposant le modèle correctement spécifié et sur base d'un échantillon d'observations des durées $T_i$ de chômage, des destinations $E_i$ de sortie et des variables explicatives $X_{it}$ pour $n$ individus tirés au hasard, estimer la "vraie valeur" inconnue $\beta^o$ du vecteur de paramètres $\beta$ par la méthode du Maximum de Vraisemblance (MV).

- La forme de l'estimateur MV (ou plus précisément MVc) de la "vraie valeur" inconnue $\beta^o$ du vecteur de paramètres $\beta$ des fonctions de hasard spécifique conditionnelles $\lambda_l(t, X^{it}; \beta_l)$ dépend du schéma d'échantillonnage sur base duquel on suppose qu'est obtenu un échantillon d'observations sur les durées $T_i$ de chômage, les destinations $E_i$ de sortie et les variables explicatives $X_{it}$, et en particulier du fait de savoir si ces durées, ces destinations et ces variables explicatives sont observées de façon complète ou non (la cas typique est le cas où les durées, les destinations et les variables explicatives ne sont pas observées de façon complète). On considère ci-dessous deux schémas d'échantillonnage : échantillonnage aléatoire simple et échantillonnage aléatoire avec censure à droite. Ces deux schémas d'échantillonnage correspondent à deux cas de figure différents concernant l'observabilité des durées $T_i$ de chômage, des destinations $E_i$ de sortie et des variables explicatives $X_{it}$. Pour chacun d'eux , on dérive l'estimateur MVc de $\beta^o$. On notera que le cas de l'échantillonnage aléatoire simple peut être regardé comme un cas particulier de l'échantillonnage aléatoire avec censure à droite. En d'autres termes, ce dernier cas couvre en fait le cas précédent. A l'instar de ce que nous avons fait pour le modèle de base, en suivant la même approche, d'autres schémas d'échantillonnage peuvent évidemment être envisagés, mais nous ne le ferons pas ici.

### A. Echantillonnage aléatoire simple

- Dans ce schéma d'échantillonnage, on suppose qu'un échantillon d'observations de taille $n$ est obtenu en tirant au hasard $n$ individus dans la population cible, et que pour chaque individu tiré, on observe de façon complète la durée $T_i$ de son épisode de chômage, sa destinations $E_i$ de sortie ainsi que l'ensemble des variables explicatives $X_{it}$ de $t = 0$ jusque $t = T_i$. Pour chaque individu tiré, on observe donc une réalisation du triplet $(T_i, E_i, X^{iT_i})$, et l'échantillon d'observations pour les $n$ individus tirés est une réalisation du $n$-uple de triplets $((T_1, E_1, X^{1T_1}), ..., (T_n, E_n, X^{nT_n}))$.

On notera que ce mode d'échantillonnage suppose que chacun des individus tirés soit suivi jusqu'à la fin de son épisode de chômage, qu'elle qu'en soit la longueur. Il s'agit évidemment d'un mode d'échantillonnage peu probable en pratique, mais qu'il est intéressant de considérer à titre de référence.

- On commence par raisonner comme si, pour chaque individu tiré, on observait une réalisation du triplet $(T_i, E_i, X^{i\infty})$, i.e. comme si les variables explicatives $X_{it}$ étaient toujours entièrement observées, de $t = 0$ jusque $t = \infty$. Dans ce cas de figure, étant donné l'hypothèse d'exogénéité stricte des $X_{it}$, si le modèle est bien correctement spécifié, la distribution jointe conditionnelle à $X^{i\infty}$ de la durée $T_i$ de chômage et de la destination $E_i$ de sortie d'un individu $i$ tiré au hasard est donnée par la fonction de densité jointe conditionnelle :

$$\begin{aligned}\overline{f}_L(t, l|X^{i\infty}) &= I\!P\left[T_i = t, E_i = l|X^{i\infty}\right] = I\!P\left[T_i = t, E_i = l|X^{it}\right] \\ &= f_L(t, l, X^{it}; \beta^o) = \lambda_l(t, X^{it}; \beta_l^o) S(t, X^{it}; \beta^o), \quad \begin{array}{l}\forall\, t = 0, 1, 2, ... \\ \forall\, l = 1, 2, ..., L\end{array}\end{aligned}$$

Pour un échantillon de $n$ individus tirés au hasard[29], les $n$ triplets $(T_1, E_1, X^{1\infty})$, ..., $(T_n, E_n, X^{n\infty})$ sont indépendants et de même distribution, et un estimateur MVc $\hat{\beta}_n$ de $\beta^o$ est dès lors donné par :

$$\begin{aligned}\hat{\beta}_n &= \text{Argmax}_{\beta\in\Theta} \sum_{i=1}^{n} \ln f_L(T_i, E_i, X^{iT_i}; \beta) \\ &= \text{Argmax}_{\beta\in\Theta} \sum_{i=1}^{n}\sum_{l=1}^{L} \mathbb{I}_{(E_i=l)} \ln \lambda_l(T_i, X^{iT_i}; \beta_l) + \sum_{i=1}^{n} \ln S(T_i, X^{iT_i}; \beta)\end{aligned}$$

où :

$$\ln S(T_i, X^{iT_i}; \beta) = \begin{cases} 0 & , \text{si } T_i = 0 \\ \sum_{t^*=0}^{T_i-1} \ln\left(1 - \sum_{l=1}^{L} \lambda_l(t^*, X^{it^*}; \beta_l)\right) & , \forall\, T_i = 1, 2, ... \end{cases}$$

- On constate que l'estimateur MVc $\hat{\beta}_n$ de $\beta^o$ ne dépend pas de l'ensemble des variables de l'échantillon $((T_1, E_1, X^{1\infty}), ..., (T_n, E_n, X^{n\infty}))$, mais, étant donné l'hypothèse d'exogénéité stricte des $X_{it}$, seulement de $((T_1, E_1, X^{1T_1}), ...,$ $(T_n, E_n, X^{nT_n}))$. Autrement dit, pour estimer $\beta^o$ par l'estimateur MVc $\hat{\beta}_n$, il n'est pas nécessaire de disposer d'observations de $((T_1, E_1, X^{1\infty}), ..., (T_n, E_n, X^{n\infty}))$, mais seulement d'observations de $((T_1, E_1, X^{1T_1}), ..., (T_n, E_n, X^{nT_n}))$. $\hat{\beta}_n$ est donc

[29] Au sens strict, avec remise.

un estimateur MVc adéquat pour estimer $\beta^o$ sur base des observations dont on dispose dans le schéma d'échantillonnage supposé.

- L'estimateur $\hat{\beta}_n$ étant un estimateur MVc standard, sous des conditions de régularité générales, il possède les propriétés habituelles des estimateurs MVc, à savoir (si, comme supposé, le modèle est bien correctement spécifié et les $X_{it}$ strictement exogènes) :

$$\hat{\beta}_n \xrightarrow{p} \beta^o \quad \text{et} \quad V^{o\frac{1}{2}}\sqrt{n}(\hat{\beta}_n - \beta^o) \xrightarrow{d} \mathcal{N}(0, I_p)$$

où[30] :

$$\begin{aligned} V^o &= \frac{1}{n}\sum_{i=1}^{n} E\left[\left(\tfrac{\partial \ln f_L(T_i,E_i,X^{iT_i};\beta^o)}{\partial\beta}\right)\left(\tfrac{\partial \ln f_L(T_i,E_i,X^{iT_i};\beta^o)}{\partial\beta}\right)'\right] \\ &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial^2 \ln f_L(T_i,E_i,X^{iT_i};\beta^o)}{\partial\beta\partial\beta'}\right] \end{aligned}$$

soit en termes d'approximation utilisable est échantillon fini :

$$\hat{\beta}_n \approx \mathcal{N}(\beta^o, \hat{V}_n^{-1}/n)$$

avec

$$\hat{V}_n = -\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2 \ln f_L(T_i,E_i,X^{iT_i};\hat{\beta}_n)}{\partial\beta\partial\beta'}$$

### B. Echantillonnage aléatoire avec censure à droite

- Dans ce schéma d'échantillonnage, on suppose qu'un échantillon de taille $n$ est obtenu en tirant au hasard $n$ individus dans la population cible, et que chaque individu tiré est observé, tant en ce qui concerne la durée $T_i$ de son épisode de chômage, sa destination $E_i$ de sa sortie et l'ensemble de ses variables explicatives $X_{it}$, au maximum seulement durant une période allant de $t = 0$ jusque $t = C_i$, où $C_i = 0, 1, 2, ...$ désigne un point de censure à droite, fixe ou aléatoire.

  Sont ainsi observés, lorsque la durée $T_i$ de l'épisode de chômage est inférieur ou égale à $C_i$, à la fois la durée $T_i$ exacte de l'épisode de chômage, sa destination $E_i$ de sa sortie et l'ensemble des variables explicatives $X_{it}$ de $t = 0$ jusque $t = T_i$, et lorsque la durée $T_i$ de l'épisode de chômage est supérieure ou égale à $C_i + 1$ (= supérieure à $C_i$), simplement le fait que l'on a $T_i \geq C_i + 1$ (sans information sur $E_i$) et l'ensemble des variables explicatives $X_{it}$ de $t = 0$ jusque $t = C_i$ (et non $C_i + 1$).

  Pour chaque individu tiré, on observe donc plus, comme dans le cas de l'échantil-

[30] On notera que la matrice de variance-covariance $V^o$ ne dépend pas de $n$. Cela résulte simplement du fait que, comme l'échantillon $((T_1, E_i, X^{1\infty}), ..., (T_n, E_i, X^{n\infty}))$ est i.i.d., on a :

$$E\left[\left(\tfrac{\partial \ln f_L(T_i,E_i,X^{iT_i};\beta^o)}{\partial\beta}\right)\left(\tfrac{\partial \ln f_L(T_i,E_i,X^{iT_i};\beta^o)}{\partial\beta}\right)'\right] = -E\left[\frac{\partial^2 \ln f(T_i,E_i,X^{iT_i};\beta^o)}{\partial\beta\partial\beta'}\right] = V^o, \quad \forall\, i = 1, ..., n$$

lonnage aléatoire simple, une réalisation du triplet $(T_i, E_i, X^{iT_i})$, mais une réalisation du quadruplet $(T_i^*, E_i^*, X^{iM_i^*}, C_i)$, où :

- $T_i^*$ désigne la durée de chômage effectivement observée de l'individu $i$ étant donné le mécanisme d'échantillonnage, durée qui est égale à :

$$T_i^* = \begin{cases} T_i & , \text{si } T_i \leq C_i \\ C_i + 1 & , \text{si } T_i \geq C_i + 1 \end{cases}$$

  où $T_i^*$ est par convention fixé à $C_i + 1$ lorsque $T_i \geq C_i + 1$ ($\Leftrightarrow T_i > C_i$), de sorte que l'ensemble $\mathbb{T}_i$ des valeurs possibles de $T_i^*$ est $\mathbb{T}_i \equiv \{0, 1, 2, ..., C_i, C_i + 1\}$.
- $E_i^*$ désigne la destination de sortie effectivement observée de l'individu $i$ étant donné le mécanisme d'échantillonnage, destination qui est égale à :

$$E_i^* = \begin{cases} E_i & , \text{si } T_i \leq C_i \\ 0 & , \text{si } T_i \geq C_i + 1 \end{cases}$$

  où $E_i^*$ est par convention fixé à 0 lorsque $T_i \geq C_i + 1$ ($\Leftrightarrow T_i > C_i$), de sorte que l'ensemble $\mathbb{E}_i$ des valeurs possibles de $E_i^*$ est $\mathbb{E}_i \equiv \{0, 1, 2, ..., L\}$.
- $M_i^* = \min(T_i^*, C_i)$, ce qui reflète le fait que, lorsque $T_i \leq C_i$ (et donc $T_i^* = T_i$), on observe les variables explicatives $X_{it}$ de $t = 0$ jusque $t = T_i^* = T_i$, et lorsque $T_i \geq C_i + 1$ (et donc $T_i^* = C_i + 1$), on observe les variables explicatives $X_{it}$ de $t = 0$ jusque $t = T_i^* - 1 = C_i$ (et non $T_i^* = C_i + 1$).

et l'échantillon d'observations pour les $n$ individus tirés est une réalisation du $n$-uple de quadruplets $((T_1^*, E_1^*, X^{1M_1^*}, C_1), ..., (T_n^*, E_n^*, X^{nM_n^*}, C_n))$.

On note $\mathbb{G}_i$ l'ensemble des valeurs possibles du couple $(T_i^*, E_i^*)$. On remarquera $\mathbb{G}_i$ n'est pas simplement égal à $\mathbb{T}_i \times \mathbb{E}_i$. Un certain nombre de couples de $\mathbb{T}_i \times \mathbb{E}_i$ sont en effet impossibles, à savoir tous les couples $(t, l)$ où $t \leq C_i$ et $l = 0$, et encore tous les couples $(t, l)$ où $t = C_i + 1$ et $l \neq 0$.

Ce mode d'échantillonnage est le mode d'échantillonnage qui est typiquement rencontré en pratique.

- On raisonne dans un premier temps comme si, pour chaque individu tiré, on observait une réalisation du quadruplet $(T_i^*, E_i^*, X^{i\infty}, C_i)$, i.e. comme si les variables explicatives $X_{it}$ étaient toujours entièrement observées, de $t = 0$ jusque $t = \infty$.

  On commence par supposer que, conditionnellement à $X^{i\infty}$, la durée $T_i$ de chômage et la destination $E_i$ de sortie (qu'elles soient (si $T_i \leq C_i$) ou non (si $T_i \geq C_i + 1$) effectivement observées de façon exacte) d'un individu $i$ pris au hasard sont indépendantes de son point de censure $C_i$, soit que l'on a :

$$I\!P\left[T_i = t, E_i = l | X^{i\infty}, C_i\right] = I\!P\left[T_i = t, E_i = l | X^{i\infty}\right], \quad \begin{array}{l} \forall\, t = 0, 1, 2, ... \\ \forall\, l = 1, 2, ..., L \end{array} \tag{30}$$

  autrement dit que, sachant $X^{i\infty}$, la connaissance du point de censure $C_i$, c'est-à-dire que l'on observe de façon complète le parcours de l'individu $i$ seulement durant la période allant de $t = 0$ jusque $t = C_i$, n'apporte aucune information sur sa durée effective (simplement non observée de façon exacte si $T_i \geq C_i + 1$) de chômage $T_i$ et sa destination de sortie $E_i$ (aussi non observée si $T_i \geq C_i + 1$). Cette hypothèse est automatiquement satisfaite si le point de censure $C_i$ est fixe, non aléatoire, ou encore si c'est une fonction de $X^{i\infty}$.

La condition (30) d'indépendance conditionnelle entre $(T_i, E_i)$ et $C_i$ implique (cf. preuve dans l'Annexe I) que :

$$I\!P\left[T_i \geq t | X^{i\infty}, C_i\right] = I\!P\left[T_i \geq t | X^{i\infty}\right], \quad \forall\, t = 0, 1, 2, ... \tag{31}$$

et par conséquent aussi que[31] :

$$\begin{aligned} & I\!P\left[T_i = t, E_i = l | T_i \geq t, X^{i\infty}, C_i\right] \\ = \; & I\!P\left[T_i = t, E_i = l | T_i \geq t, X^{i\infty}\right], \quad \begin{matrix} \forall\, t = 0, 1, 2, ... \\ \forall\, l = 1, 2, ..., L \end{matrix} \end{aligned}$$

Etant donné la condition (30) d'indépendance conditionnelle entre $(T_i, E_i)$ et $C_i$, et son implication (31), on peut déduire la distribution conditionnelle à $(X^{i\infty}, C_i)$ de la durée de chômage et de la destination de sortie effectivement observées $(T_i^*, E_i^*)$ d'un individu $i$ pris au hasard. Elle est donnée par la fonction de densité jointe conditionnelle :

$$\overline{f}_L^*(t, l | X^{i\infty}, C_i) = I\!P\left[T_i^* = t, E_i^* = l | X^{i\infty}, C_i\right], \quad \forall\, (t, l) \in \mathbb{G}_i$$

$$= \begin{cases} I\!P[T_i = t, E_i = l | X^{i\infty}, C_i] = I\!P[T_i = t, E_i = l | X^{i\infty}] & , \text{ si } t \leq C_i \ (\Rightarrow l \neq 0) \\ I\!P[T_i \geq t | X^{i\infty}, C_i] = I\!P[T_i \geq t | X^{i\infty}] & , \text{ si } t = C_i + 1 \ (\Rightarrow l = 0) \end{cases}$$

qui, puisqu'étant donné l'hypothèse d'exogénéité stricte des $X_{it}$, si le modèle est bien correctement spécifié, on a :

$$\begin{aligned} I\!P\left[T_i = t, E_i = l | X^{i\infty}\right] &= I\!P\left[T_i = t, E_i = l | X^{it}\right] = f_L(t, l, X^{it}; \beta^o) \\ &= \lambda_l(t, X^{it}; \beta_l^o) S(t, X^{it}; \beta^o), \quad \begin{matrix} \forall\, t = 0, 1, 2, ... \\ \forall\, l = 1, 2, ..., L \end{matrix} \end{aligned}$$

et

$$\begin{aligned} I\!P\left[T_i \geq t | X^{i\infty}\right] &= I\!P\left[T_i \geq t | X^{it}\right] \\ &= S(t, X^{it}; \beta^o), \quad \forall\, t = 0, 1, 2, ... \end{aligned}$$

[31] En effet, étant donné (30) et (31) et le fait que $(T_i = t) \Rightarrow (T_i \geq t)$, $\forall\, t = 0, 1, 2, ..., \forall\, l = 1, 2, ..., L$, on a :

$$\begin{aligned} I\!P\left[T_i = t, E_i = l | T_i \geq t, X^{i\infty}, C_i\right] &= \frac{I\!P\left[T_i = t, E_i = l | X^{i\infty}, C_i\right]}{I\!P\left[T_i \geq t | X^{i\infty}, C_i\right]} = \frac{I\!P\left[T_i = t, E_i = l | X^{i\infty}\right]}{I\!P\left[T_i \geq t | X^{i\infty}\right]} \\ &= I\!P\left[T_i = t, E_i = l | T_i \geq t, X^{i\infty}\right] \end{aligned}$$

et que, étant donné (27)[32], on a également :

$$\begin{aligned} S(t, X^{it}; \beta^o) &= I\!P\left[T_i \geq t | X^{it}\right] = I\!P\left[T_i \geq t | X^{it-1}\right] = S(t, X^{it-1}; \beta^o) \\ &= \prod_{t^*=0}^{t-1} \left(1 - \sum_{l=1}^{L} \lambda_l(t^*, X^{it^*}; \beta_l^o)\right), \ \forall\, t = 1, 2, ... \end{aligned} \tag{32}$$

est égale à :

$$\begin{aligned} & \overline{f}^*_L(t, l | X^{i\infty}, C_i) = I\!P\left[T_i^* = t, E_i^* = l | X^{i\infty}, C_i\right], \quad \forall\, (t, l) \in \mathbb{G}_i \\ = \ & I\!P\left[T_i^* = t, E_i^* = l | X^{im_i^*}, C_i\right] \\ = \ & f_L^*(t, l, X^{im_i^*}, C_i; \beta^o) = \begin{cases} \lambda_l(t, X^{it}; \beta_l^o) S(t, X^{it}; \beta^o) & \text{, si } t \leq C_i \\ S(t, X^{it-1}; \beta^o) & \text{, si } t = C_i + 1 \end{cases} \end{aligned}$$

où $m_i^* = \min(t, C_i)$.

Pour un échantillon de $n$ individus tirés au hasard[33], les $n$ quadruplets $(T_1^*, E_1^*, X^{n\infty}, C_i), ..., (T_n^*, E_n^*, X^{n\infty}, C_n)$ sont indépendants et de même distribution, et un estimateur MVc $\hat{\beta}_n$ de $\beta^o$ est dès lors donné par :

$$\begin{aligned} \hat{\beta}_n &= \text{Argmax}_{\beta \in \Theta} \sum_{i=1}^{n} \ln f_L^*(T_i^*, E_i^*, X^{iM_i^*}, C_i; \beta) \\ &= \text{Argmax}_{\beta \in \Theta} \sum_{i=1}^{n} \left[ \mathbb{I}_{(T_i^* \leq C_i)} \left( \sum_{l=1}^{L} \mathbb{I}_{(E_i^* = l)} \ln \lambda_l(T_i^*, X^{iT_i^*}, \beta_l) + \ln S(T_i^*, X^{iT_i^*}, \beta) \right) \right. \\ & \qquad \left. + \, \mathbb{I}_{(T_i^* = C_i + 1)} \ln S(T_i^*, X^{iT_i^* - 1}, \beta) \right] \end{aligned}$$

qu'on peut encore écrire :

$$\begin{aligned} \hat{\beta}_n &= \text{Argmax}_{\beta \in \Theta} \sum_{i=1}^{n} \ln f_L^*(T_i^*, E_i^*, X^{iM_i^*}, C_i; \beta) \\ &= \text{Argmax}_{\beta \in \Theta} \sum_{i=1}^{n} \sum_{l=1}^{L} \mathbb{I}_{(T_i^* \leq C_i \text{ et } E_i^* = l)} \ln \lambda_l(T_i^*, X^{iT_i^*}, \beta_l) + \sum_{i=1}^{n} \ln S(T_i^*, X^{iM_i^*}, \beta) \end{aligned}$$

où $M_i^* = \min(T_i^*, C_i)$, les $\mathbb{I}_{(.)}$ sont des variables indicatrices et :

$$\ln S(T_i^*, X^{iM_i^*}, \beta) = \begin{cases} 0 & \text{, si } T_i^* = 0 \\ \sum_{t^*=0}^{T_i^*-1} \ln \left(1 - \sum_{l=1}^{L} \lambda_l(t^*, X^{it^*}; \beta_l)\right) & \text{, } \forall\, T_i^* = 1, 2, ..., C_i + 1 \end{cases}$$

- De façon semblable au cas de l'échantillonnage aléatoire simple, on constate que l'estimateur MVc $\hat{\beta}_n$ de $\beta^o$ ne dépend pas de l'ensemble des variables de l'échantillon $((T_1^*, E_1^*, X^{1\infty}, C_1), ..., (T_n^*, E_n^*, X^{n\infty}, C_n))$, mais, étant donné l'hypo-

[32] C'est-à-dire que la fonction de survie $S(t, X^{it})$ ne dépend en fait que de $X^{it-1}$.

[33] Au sens strict, avec remise.

thèse d'exogénéité stricte des $X_{it}$ (et (32)), seulement de $((T_1^*, E_1^*, X^{1M_1^*}, C_1), ..., (T_n^*, E_n^*, X^{nM_n^*}, C_n))$. Autrement dit, pour estimer $\beta^o$ par l'estimateur MVc $\hat{\beta}_n$, il n'est pas nécessaire de disposer d'observations de $((T_1^*, E_1^*, X^{1\infty}, C_1), ..., (T_n^*, E_n^*, X^{n\infty}, C_n))$, mais seulement d'observations de $((T_1^*, E_1^*, X^{1M_1^*}, C_1), ..., (T_n^*, E_n^*, X^{nM_n^*}, C_n))$. $\hat{\beta}_n$ est donc un estimateur MVc adéquat pour estimer $\beta^o$ sur base des observations dont on dispose dans le schéma d'échantillonnage supposé.

- L'estimateur $\hat{\beta}_n$ étant un estimateur MVc standard, sous des conditions de régularité générales, il possède les propriétés habituelles des estimateurs MVc, à savoir (si, comme supposé, le modèle est bien correctement spécifié et les $X_{it}$ strictement exogènes) :

$$\hat{\beta}_n \xrightarrow{p} \beta^o \quad \text{et} \quad V^{o\frac{1}{2}} \sqrt{n}(\hat{\beta}_n - \beta^o) \xrightarrow{d} \mathcal{N}(0, I_p)$$

où[34] :

$$\begin{aligned} V^o &= \frac{1}{n}\sum_{i=1}^n E\left[\left(\frac{\partial \ln f_L^*(T_i^*, E_i^*, X^{iM_i^*}, C_i; \beta^o)}{\partial \beta}\right)\left(\frac{\partial \ln f_L^*(T_i^*, E_i^*, X^{iM_i^*}, C_i; \beta^o)}{\partial \beta}\right)'\right] \\ &= -\frac{1}{n}\sum_{i=1}^n E\left[\frac{\partial^2 \ln f_L^*(T_i^*, E_i^*, X^{iM_i^*}, C_i; \beta^o)}{\partial \beta \partial \beta'}\right] \end{aligned}$$

soit en termes d'approximation utilisable est échantillon fini :

$$\hat{\beta}_n \approx \mathcal{N}(\beta^o, \hat{V}_n^{-1}/n)$$

avec

$$\hat{V}_n = -\frac{1}{n}\sum_{i=1}^n \frac{\partial^2 \ln f_L^*(T_i^*, E_i^*, X^{iM_i^*}, C_i; \hat{\beta}_n)}{\partial \beta \partial \beta'}$$

### 2.3.2. Estimation sans l'hypothèse d'exogénéité stricte des $X_{it}$

#### A. Echantillonnage aléatoire simple

- En principe, l'estimateur standard (pour $X_{it}$ exogène) reste adéquat : il est toujours convergent et asymptotiquement normal (avec la même matrice de variance-covariance asymptotique), mais correspond maintenant à un estimateur MVc partiel (et non MVc standard).

---

[34] A nouveau, on notera que la matrice de variance-covariance $V^o$ ne dépend pas de $n$. Cela résulte simplement du fait que, comme l'échantillon $((T_1^*, E_1^*, X^{1\infty}, C_1), ..., (T_n^*, E_n^*, X^{n\infty}, C_n))$ est i.i.d., on a :

$$\begin{aligned} & E\left[\left(\frac{\partial \ln f_L^*(T_i^*, E_i^*, X^{iM_i^*}, C_i; \beta^o)}{\partial \beta}\right)\left(\frac{\partial \ln f_L^*(T_i^*, E_i^*, X^{iM_i^*}, C_i; \beta^o)}{\partial \beta}\right)'\right] \\ = & -E\left[\frac{\partial^2 \ln f_L^*(T_i^*, E_i^*, X^{iM_i^*}, C_i; \beta^o)}{\partial \beta \partial \beta'}\right] = V^o, \ \forall i = 1, ..., n \end{aligned}$$

**B. Echantillonnage aléatoire avec censure à droite**

- En principe, l'estimateur standard (pour $X_{it}$ exogène) reste à nouveau adéquat : il est toujours convergent et asymptotiquement normal (avec la même matrice de variance-covariance asymptotique), mais correspond maintenant à un estimateur MVc partiel (et non MVc standard).

# 3. Annexes

# Annexe A

- Preuve de l'équivalence des conditions d'exogénéité stricte :

$$D\left(X^{\underline{it+1}}|T_i = t-k, X^{it}\right) = D\left(X^{\underline{it+1}}|X^{it}\right), \ \forall\, t = 0,1,2,..., \ \forall\, k = 0,1,2,...,t \tag{A-1}$$

et

$$D\left(X^{\underline{it+1}}|T_i = t, X^{it}\right) = D\left(X^{\underline{it+1}}|X^{it}\right), \ \forall\, t = 0,1,2,... \tag{A-2}$$

On doit montrer que (A-2) est une condition nécessaire et suffisante pour (A-1).

- Preuve que (A-2) est une condition nécessaire pour (A-1), i.e. que (A-1) $\Rightarrow$ (A-2).

  La preuve est immédiate en prenant $k = 0$ dans (A-1).

- Preuve que (A-2) est une condition suffisante pour (A-1), i.e. que (A-2) $\Rightarrow$ (A-1).

  La condition (A-2) est identique à la condition (A-1) $\forall\, t = 0,1,2,...$ et $k = 0$. Il suffit donc de montrer que (A-2) $\Rightarrow$ (A-1) $\forall\, t = 0,1,2,...$ et $\forall\, k = 1,2,...,t$.
  De (A-2), $\forall\, t = 0,1,2,...,\forall\, k = 1,2,...,t$, on a :

$$D\left(X^{\underline{it-k+1}}|T_i = t-k, X^{it-k}\right) = D\left(X^{\underline{it-k+1}}|X^{it-k}\right) \tag{A-3}$$

  où $X^{\underline{it-k+1}} = (X'_{it-k+1}, X'_{it-k+2}, ..., X'_{i\infty})'$ et $X^{it-k} = (X'_{i1}, X'_{i2}, ..., X'_{it-k})'$.

  La condition (A-3) signifie que, conditionnellement à $X^{it-k}$, $X^{\underline{it-k+1}}$ et $(T_i = t-k)$ sont indépendants :

$$X^{\underline{it-k+1}} \perp (T_i = t-k) \mid X^{it-k} \ , \ \forall\, t = 0,1,2,...,\forall\, k = 1,2,...,t \tag{A-4}$$

  La propriété (A-4) implique que, $\forall\, t = 0,1,2,...,\forall\, k = 1,2,...,t$,on a :

$$\begin{aligned} D\left(T_i = t-k|X^{it-k}, X^{\underline{it-k+1}}\right) &= D\left(T_i = t-k|X^{it-k}, X^{\underline{i(t-k+1,t)}}, X^{\underline{it+1}}\right) \\ &= D\left(T_i = t-k|X^{it-k}\right) \end{aligned}$$

  où $D(.|.)$ représente des distributions conditionnelles et $X^{\underline{i(t-k+1,t)}} = (X'_{i(t-k+1)}, ...., X'_{it})'$, et donc aussi que, comme $X^{it} = (X^{it-k\prime}, X^{\underline{i(t-k+1,t)}\prime})'$,

$\forall\, t = 0, 1, 2, ..., \forall\, k = 1, 2, ..., t$, on a :

$$\begin{aligned} & D\left(T_i = t - k | X^{it-k}, X^{\underline{it-k+1}}\right) \\ = \; & D\left(T_i = t - k | X^{it-k}, X^{\underline{i(t-k+1,t)}}, X^{\underline{it+1}}\right) \\ = \; & D\left(T_i = t - k | X^{it}, X^{\underline{it+1}}\right) = D\left(T_i = t - k | X^{it}\right) \end{aligned} \quad \text{(A-5)}$$

L'égalité de (A-5) signifie que, conditionnellement à $X^{it}$, $X^{\underline{it+1}}$ et $(T_i = t - k)$ sont indépendants :

$$X^{\underline{it+1}} \perp (T_i = t - k) \mid X^{it}\ , \;\; \forall\, t = 0, 1, 2, ..., \forall\, k = 1, 2, ..., t$$

ce qui implique que :

$$D\left(X^{\underline{it+1}} | T_i = t - k, X^{it}\right) = D\left(X^{\underline{it+1}} | X^{it}\right), \;\; \forall\, t = 0, 1, 2, ..., \forall\, k = 1, 2, ..., t$$

CQFD.

# Annexe B

• Preuve que la condition d'exogénéité stricte :

$$D\left(X^{\underline{it+1}} | T_i = t - k, X^{it}\right) = D\left(X^{\underline{it+1}} | X^{it}\right), \;\; \forall t = 0, 1, 2, ..., \;\; \forall k = 0, 1, 2, ..., t \quad \text{(B-1)}$$

implique :

$$I\!P\left[T_i = t | X^{it}, X^{\underline{it+1}}\right] = I\!P\left[T_i = t | X^{it}\right], \;\; \forall\, t = 0, 1, 2, ... \quad \text{(B-2)}$$

$$I\!P\left[T_i \geq t | X^{it}, X^{\underline{it+1}}\right] = I\!P\left[T_i \geq t | X^{it}\right], \;\; \forall\, t = 0, 1, 2, ... \quad \text{(B-3)}$$

et

$$I\!P\left[T_i = t | T_i \geq t, X^{it}, X^{\underline{it+1}}\right] = I\!P\left[T_i = t | T_i \geq t, X^{it}\right], \;\; \forall\, t = 0, 1, 2, ... \quad \text{(B-4)}$$

• Preuve que (B-1) implique (B-2).

La condition (B-1) signifie que, conditionnellement à $X^{it}$, $X^{\underline{it+1}}$ et $(T_i = t - k)$ sont indépendants :

$$X^{\underline{it+1}} \perp (T_i = t - k) \mid X^{it}\ , \;\; \forall t = 0, 1, 2, ..., \;\; \forall k = 0, 1, 2, ..., t$$

Cela implique que, $\forall\, t = 0, 1, 2, ..., \forall\, k = 0, 1, 2, ..., t$, on a :

$$I\!P\left[T_i = t - k | X^{it}, X^{\underline{it+1}}\right] = I\!P\left[T_i = t - k | X^{it}\right] \quad \text{(B-5)}$$

soit, en prenant $k = 0$ :

$$I\!P\left[T_i = t | X^{it}, X^{\underline{it+1}}\right] = I\!P\left[T_i = t | X^{it}\right]$$

CQFD.

- Preuve que (B-1) implique (B-3).

  $\forall\, t = 0, 1, 2, ...$, on a :

$$\begin{aligned}
I\!\!P\left[T_i \geq t | X^{it}, X^{\underline{it+1}}\right] &= 1 - I\!\!P\left[T_i \leq t-1 | X^{it}, X^{\underline{it+1}}\right] \\
&= 1 - \sum_{t^*=0}^{t-1} I\!\!P\left[T_i = t^* | X^{it}, X^{\underline{it+1}}\right]
\end{aligned}$$

  soit, étant donné (B-5),

$$\begin{aligned}
I\!\!P\left[T_i \geq t | X^{it}, X^{\underline{it+1}}\right] &= 1 - \sum_{t^*=0}^{t-1} I\!\!P\left[T_i = t^* | X^{it}\right] \\
&= 1 - I\!\!P\left[T_i \leq t-1 | X^{it}\right] \\
&= I\!\!P\left[T_i \geq t | X^{it}\right] \qquad \text{(B-6)}
\end{aligned}$$

CQFD.

- Preuve que (B-1) implique (B-4).

  $\forall\, t = 0, 1, 2, ...$, on a :

$$I\!\!P\left[T_i = t | T_i \geq t, X^{it}, X^{\underline{it+1}}\right] = \frac{I\!\!P[T_i = t, T_i \geq t | X^{it}, X^{\underline{it+1}}]}{I\!\!P[T_i \geq t | X^{it}, X^{\underline{it+1}}]}, \ \ \forall\, t = 0, 1, 2, ...$$

  soit, comme $(T_i = t)$ implique $(T_i \geq t)$,

$$I\!\!P\left[T_i = t | T_i \geq t, X^{it}, X^{\underline{it+1}}\right] = \frac{I\!\!P[T_i = t | X^{it}, X^{\underline{it+1}}]}{I\!\!P[T_i \geq t | X^{it}, X^{\underline{it+1}}]}$$

  soit encore, étant donné (B-5), (B-6) et l'implication évoquée ci-dessus,

$$\begin{aligned}
I\!\!P\left[T_i = t | T_i \geq t, X^{it}, X^{\underline{it+1}}\right] &= \frac{I\!\!P[T_i = t | X^{it}]}{I\!\!P[T_i \geq t | X^{it}]} \\
&= \frac{I\!\!P[T_i = t, T_i \geq t | X^{it}]}{I\!\!P[T_i \geq t | X^{it}]} \\
&= I\!\!P\left[T_i = t | T_i \geq t, X^{it}\right]
\end{aligned}$$

CQFD.

# Annexe C

- Preuve directe que, sous l'hypothèse d'exogénéité stricte des $X_{it}$, qui implique que

$$I\!\!P\left[T_i = t | X^{i\infty}\right] = I\!\!P\left[T_i = t | X^{it}\right], \ \ \forall\, t = 0, 1, 2, ... \qquad \text{(C-1)}$$

  on a bien

$$I\!\!P\left[T_i \geq t | X^{it}\right] = I\!\!P\left[T_i \geq t | X^{it-1}\right] \ \ \forall\, t = 1, 2, ...$$

où $X^{it-1} = \left(X'_{i0}, X'_{i1}, ..., X'_{it-1}\right)'$.

$\forall\, t = 1, 2, ...$, on a :

$$\begin{aligned} I\!P\left[T_i \geq t | X^{it}\right] &= 1 - I\!P\left[T_i \leq t-1 | X^{it}\right] \\ &= 1 - \sum_{t^*=0}^{t-1} I\!P\left[T_i = t^* | X^{it}\right] \end{aligned}$$

soit, étant donné (C-1) et que $X^{it} = (X^{it-1\prime}, X'_{it})'$,

$$\begin{aligned} I\!P\left[T_i \geq t | X^{it}\right] &= 1 - \sum_{t^*=0}^{t-1} I\!P\left[T_i = t^* | X^{it-1}\right] \\ &= 1 - I\!P\left[T_i \leq t-1 | X^{it-1}\right] \\ &= I\!P\left[T_i \geq t | X^{it-1}\right] \end{aligned}$$

CQFD.

# Annexe D

- Calcul de la fonction de survie conditionnelle $S(t, X^{it})$ dans le cas où l'hypothèse d'exogénéité stricte des $X_{it}$ ne tient pas.

- En vue du calcul de la fonction de survie conditionnelle $S(t, X^{it})$, on considère tout d'abord la probabilité jointe (en supposant $X_{it}$ discret et par abus de notation) $I\!P[T_i \geq t, X^{it}]$. Par conditionnement successifs, on a[35] :

$$\begin{aligned} I\!P\left[T_i \geq t, X^{it}\right] &= I\!P\left[T_i \geq t, X_{it} | T_i \geq t-1, X^{it-1}\right] I\!P\left[T_i \geq t-1, X^{it-1}\right] \\ &= I\!P\left[T_i \geq t, X_{it} | T_i \geq t-1, X^{it-1}\right] \\ &\quad \times I\!P\left[T_i \geq t-1, X_{it-1} | T_i \geq t-2, X^{it-2}\right] I\!P\left[T_i \geq t-2, X^{it-2}\right] \\ &= I\!P\left[T_i \geq t, X_{it} | T_i \geq t-1, X^{it-1}\right] \\ &\quad \times I\!P\left[T_i \geq t-1, X_{it-1} | T_i \geq t-2, X^{it-2}\right] \\ &\quad \times I\!P\left[T_i \geq t-2, X_{it-2} | T_i \geq t-3, X^{it-3}\right] I\!P\left[T_i \geq t-3, X^{it-3}\right] \\ &= \text{etc...} \end{aligned}$$

[35] La validité de la relation récurrente :

$$I\!P\left[T_i \geq t, X^{it}\right] = I\!P\left[T_i \geq t, X_{it} | T_i \geq t-1, X^{it-1}\right] I\!P\left[T_i \geq t-1, X^{it-1}\right]$$

vient d'une part du fait que $X^{it} = (X^{it-1\prime}, X'_{it})'$, et d'autre part du fait que, façon semblable au cas de la note 3, l'événement $(T_i \geq t)$ implique l'événement $(T_i \geq t-1)$, de sorte qu'on a :

$$I\!P\left[T_i \geq t, X^{it}\right] = I\!P\left[T_i \geq t, T_i \geq t-1, X_{it}, X^{it-1}\right] = I\!P\left[T_i \geq t, X_{it}, T_i \geq t-1, X^{it-1}\right]$$

et donc, par conditionnement :

$$I\!P\left[T_i \geq t, X^{it}\right] = I\!P\left[T_i \geq t, X_{it} | T_i \geq t-1, X^{it-1}\right] I\!P\left[T_i \geq t-1, X^{it-1}\right]$$

En décomposant ainsi $I\!\!P[T_i \geq t, X^{it}]$ jusqu'à $t = 0$, $\forall\, t = 1, 2, ...$, on a :

$$\begin{aligned} I\!\!P\left[T_i \geq t, X^{it}\right] &= I\!\!P\left[T_i \geq 0, X^{i0}\right] \prod_{t^*=1}^{t} I\!\!P\left[T_i \geq t^*, X_{it^*} | T_i \geq t^* - 1, X^{it^*-1}\right] \\ &= I\!\!P\left(X^{i0}\right) \prod_{t^*=1}^{t} I\!\!P\left[T_i \geq t^*, X_{it^*} | T_i \geq t^* - 1, X^{it^*-1}\right] \qquad \text{(D-1)} \end{aligned}$$

puisque pour $t = 0$, on a simplement :

$$I\!\!P\left[T_i \geq 0, X^{i0}\right] = I\!\!P\left(X^{i0}\right) \quad (\text{car } (T_i \geq 0) \text{ est certain}) \qquad \text{(D-2)}$$

Par conditionnement supplémentaire, on a encore :

$$\begin{aligned} & I\!\!P\left[T_i \geq t^*, X_{it^*} | T_i \geq t^* - 1, X^{it^*-1}\right] \\ = \; & I\!\!P\left[X_{it^*} | T_i \geq t^*, T_i \geq t^* - 1, X^{it^*-1}\right] I\!\!P\left[T_i \geq t^* | T_i \geq t^* - 1, X^{it^*-1}\right] \end{aligned}$$

soit, comme

$$\begin{aligned} & I\!\!P\left[X_{it^*} | T_i \geq t^*, T_i \geq t^* - 1, X^{it^*-1}\right] \\ = \; & I\!\!P\left[X_{it^*} | T_i \geq t^*, X^{it^*-1}\right] \quad (\text{car } \; (T_i \geq t^*) \Rightarrow (T_i \geq t^* - 1)) \end{aligned}$$

et

$$\begin{aligned} & I\!\!P\left[T_i \geq t^* | T_i \geq t^* - 1, X^{it^*-1}\right] \\ = \; & 1 - I\!\!P\left[T_i \leq t^* - 1 | T_i \geq t^* - 1, X^{it^*-1}\right] \\ = \; & 1 - I\!\!P\left[T_i = t^* - 1 | T_i \geq t^* - 1, X^{it^*-1}\right] \\ = \; & 1 - \lambda(t^* - 1, X^{it^*-1}) \end{aligned}$$

où $\lambda(.,.)$ est la fonction de hasard définie par l'équation (1),

$$\begin{aligned} & I\!\!P\left[T_i \geq t^*, X_{it^*} | T_i \geq t^* - 1, X^{it^*-1}\right] \\ = \; & \left(1 - \lambda(t^* - 1, X^{it^*-1})\right) I\!\!P\left[X_{it^*} | T_i \geq t^*, X^{it^*-1}\right] \qquad \text{(D-3)} \end{aligned}$$

En utilisant (D-2) et (D-3), on déduit finalement que, $\forall\, t = 1, 2, ...$, on a :

$$\begin{aligned} I\!\!P\left[T_i \geq t, X^{it}\right] &= I\!\!P\left[T_i \geq 0, X^{i0}\right] \prod_{t^*=1}^{t} I\!\!P\left[T_i \geq t^*, X_{it^*} | T_i \geq t^* - 1, X^{it^*-1}\right] \\ &= I\!\!P\left(X^{i0}\right) \prod_{t^*=1}^{t} \left(1 - \lambda(t^* - 1, X^{it^*-1})\right) I\!\!P\left[X_{it^*} | T_i \geq t^*, X^{it^*-1}\right] \\ &= I\!\!P\left(X^{i0}\right) \prod_{t^*=0}^{t-1} \left(1 - \lambda(t^*, X^{it^*})\right) I\!\!P\left[X_{it^*+1} | T_i \geq t^* + 1, X^{it^*}\right] \\ &= \prod_{t^*=0}^{t-1} \left(1 - \lambda(t^*, X^{it^*})\right) \times \prod_{t^*=0}^{t-1} I\!\!P\left[X_{it^*+1} | T_i \geq t^* + 1, X^{it^*}\right] \times I\!\!P\left(X^{i0}\right) \end{aligned}$$

tandis que pour $t = 0$, on a simplement :

$$I\!P\left[T_i \geq t, X^{it}\right] = I\!P\left(X^{i0}\right)$$

- Ayant obtenu la probabilité jointe $I\!P[T_i \geq t, X^{it}]$, on déduit aisément la fonction de survie conditionnelle $S(t, X^{it})$ (par commodité, on continue à supposer $X_{it}$ discret et à abuser des notations).

  Par définition, on a :

$$S(t, X^{it}) = I\!P\left[T_i \geq t | X^{it}\right] = \frac{I\!P[T_i \geq t, X^{it}]}{I\!P(X^{it})}, \;\; \forall\, t = 0, 1, 2, ...$$

  Or, par conditionnement successifs, $\forall\, t = 1, 2, ...$, on a[36] :

$$\begin{aligned} I\!P\left(X^{it}\right) &= I\!P\left[X_{it}|, X^{it-1}\right] I\!P\left(X^{it-1}\right) \\ &= I\!P\left[X_{it}|, X^{it-1}\right] I\!P\left[X_{it-1}|, X^{it-2}\right] I\!P\left(X^{it-1}\right) \\ &\;\;\vdots \\ &= I\!P\left[X_{it}|, X^{it-1}\right] I\!P\left[X_{it-1}|, X^{it-2}\right] I\!P\left[X_{it-2}|, X^{it-3}\right] \\ &\quad \times ... \times I\!P\left[X_{i1}|, X^{i0}\right] I\!P\left(X^{i0}\right) \\ &= I\!P\left(X^{i0}\right) \prod_{t^*=0}^{t-1} I\!P\left[X_{it^*+1}|X^{it^*}\right] \end{aligned}$$

  tandis qu'en $t = 0$, on a simplement $I\!P(X^{i0})$.

  En combinant les résultats obtenus pour $I\!P[T_i \geq t|X^{it}]$ et $I\!P(X^{it})$, on obtient finalement :

$$S(t, X^{it}) = I\!P\left[T_i \geq t|X^{it}\right] = \begin{cases} 1 & , \text{si } t = 0 \\ \displaystyle\prod_{t^*=0}^{t-1} \left(1 - \lambda(t^*, X^{it^*})\right) & \\ \displaystyle\times \prod_{t^*=0}^{t-1} \frac{I\!P\left[X_{it^*+1}|T_i \geq t^* + 1, X^{it^*}\right]}{I\!P[X_{it^*+1}|X^{it^*}]} & , \forall\, t = 1, 2, ... \end{cases}$$

  Cette expression générale de $S(t, X^{it})$ est différente de celle obtenue sous l'hypothèse d'exogénéité stricte des $X_{it}$, mais s'y réduit si on a :

$$I\!P\left[X_{it+1}|T_i \geq t + 1, X^{it}\right] = I\!P\left[X_{it+1}|X^{it}\right], \forall\, t = 0, 1, 2, ...$$

  On montre ci-dessous que cette condition est bien remplie sous l'hypothèse d'exogénéité stricte des $X_{it}$.

[36] Raisonnement semblable à celui de la note 35.

- Preuve que la condition :

$$I\!P\left[X_{it+1}|T_i \geq t+1, X^{it}\right] = I\!P\left[X_{it+1}|X^{it}\right], \forall\, t = 0,1,2,... \tag{D-4}$$

tient sous l'hypothèse d'exogénéité stricte des $X_{it}$ :

$$D\left(X^{\underline{it+1}}|T_i = t-k, X^{it}\right) = D\left(X^{\underline{it+1}}|X^{it}\right),\ \ \forall t = 0,1,2,...,\ \forall k = 0,1,2,...,t \tag{D-5}$$

L'hypothèse d'exogénéité stricte (D-5) signifie que, conditionnellement à $X^{it}$, $X^{\underline{it+1}}$ et $(T_i = t-k)$ sont indépendants :

$$X^{\underline{it+1}} \perp (T_i = t-k) \mid X^{it}\ ,\ \ \forall\, t = 0,1,2,...,\forall\, k = 0,1,2,...,t$$

Cette condition implique que[37] :

$$X^{\underline{it+1}} \perp (T_i = 1 \text{ ou } T_i = 2 \text{ ou } ... \text{ ou } T_i = t) \mid X^{it},\ \ \forall\, t = 0,1,2,...$$

condition qui est équivalente[38] à :

$$X^{\underline{it+1}} \perp \overline{(T_i = 1 \text{ ou } T_i = 2 \text{ ou } ... \text{ ou } T_i = t)} \mid X^{it},\ \ \forall\, t = 0,1,2,...$$

Or, l'événement contraire de $(T_i = 0$ ou $T_i = 1$ ou ... ou $T_i = t)$ est l'événement $T_i \geq t+1$. On a donc aussi :

$$X^{\underline{it+1}} \perp (T_i \geq t+1) \mid X^{it},\ \ \forall\, t = 0,1,2,...$$

ce qui implique en particulier que (car $X_{it+1} \subset X^{\underline{it+1}}$) :

$$X_{it+1} \perp (T_i \geq t+1) \mid X^{it},\ \ \forall\, t = 0,1,2,... \qquad (\text{car } X_{it+1} \subset X^{\underline{it+1}})$$

soit, en supposant $X_{it}$ discret et en abusant des notations, que :

$$I\!P\left[X_{it+1}|T_i \geq t+1, X^{it}\right] = I\!P\left[X_{it+1}|X^{it}\right], \forall\, t = 0,1,2,...$$

CQFD.

- Remarque :
  Comme nous venons de le montrer, l'hypothèse d'exogénéité stricte (D-5) implique la condition (D-4). Le fait de savoir si le contraire est aussi vrai, et donc si les conditions (D-4) et (D-5) sont en fait équivalentes, est encore à vérifier.

---

[37] Cette implication découle de la propriété de base suivante. Soit $A$ un événement quelconque et $r$ événements $B_1, B_2, ..., B_r$ mutuellement exclusifs. Si on a : $A \perp B_i,\ \forall\, i = 0,1,...,r$ ($\equiv$ prop. A), alors : $A \perp (B_1$ ou $B_2$ ou ... ou $B_r)$ ($\equiv$ prop. B).
Preuve : la prop. B tient ssi $I\!P[A \text{ et } (B_1 \text{ ou } B_2 \text{ ou } ... \text{ ou } B_r)] = I\!P(A)I\!P(B_1 \text{ ou } B_2 \text{ ou } ... \text{ ou } B_r) = I\!P(A)\left(\sum_{i=1}^r I\!P(B_i)\right)$ (car les $B_i$ sonts mutuellement exclusifs). On a $I\!P[A \text{ et } (B_1 \text{ ou } B_2 \text{ ou } ... \text{ ou } B_r)] = I\!P[(A \text{ et } B_1) \text{ ou } (A \text{ et } B_2) \text{ ou } ... \text{ ou } (A \text{ et } B_r)] = \sum_{i=1}^r I\!P(A \text{ et } B_i)$ (car les $B_i$ sonts mutuellement exclusifs). Or, si la prop A. tient, on a $I\!P(A \text{ et } B_i) = I\!P(A)I\!P(B_i),\ \forall\, i = 0,1,...,r$. On en déduit que si la prop A. tient, on a $\sum_{i=1}^r I\!P(A \text{ et } B_i) = \sum_{i=1}^r I\!P(A)I\!P(B_i) = I\!P(A)\left(\sum_{i=1}^r I\!P(B_i)\right)$, CQFD.

[38] Cette équivalence découle de la propriété standard suivante. Soit $A, B$ deux événements quelconques. On a : $A \perp B \Leftrightarrow A \perp \overline{B} \Leftrightarrow \overline{A} \perp B \Leftrightarrow \overline{A} \perp \overline{B}$.

# Annexe E

- Preuve que la condition d'indépendance conditionnelle :

$$I\!P\left[T_i = t|X^{i\infty}, C_i\right] = I\!P\left[T_i = t|X^{i\infty}\right], \quad \forall\, t = 0, 1, 2, ... \tag{E-1}$$

implique :

$$I\!P\left[T_i \geq t|X^{i\infty}, C_i\right] = I\!P\left[T_i \geq t|X^{i\infty}\right], \quad \forall\, t = 0, 1, 2, ... \tag{E-2}$$

$\forall\, t = 0, 1, 2, ...$, on a :

$$\begin{aligned} I\!P\left[T_i \geq t|X^{i\infty}, C_i\right] &= 1 - I\!P\left[T_i \leq t-1|X^{i\infty}, C_i\right] \\ &= 1 - \sum_{t^*=0}^{t-1} I\!P\left[T_i = t^*|X^{i\infty}, C_i\right] \end{aligned}$$

soit, étant donné (E-1),

$$\begin{aligned} I\!P\left[T_i \geq t|X^{it}, C_i\right] &= 1 - \sum_{t^*=0}^{t-1} I\!P\left[T_i = t^*|X^{i\infty}\right] \\ &= 1 - I\!P\left[T_i \leq t-1|X^{i\infty}\right] \\ &= I\!P\left[T_i \geq t|X^{i\infty}\right] \end{aligned}$$

CQFD.

# Annexe F

- Preuve que la condition d'indépendance conditionnelle :

$$I\!P\left[T_i = t|X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] = I\!P\left[T_i = t|X^{i\infty}\right], \quad \forall\, t = 0, 1, 2, ... \tag{F-1}$$

implique :

$$\begin{aligned} & I\!P\left[t_1 \leq T_i \leq t_2|X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] \\ = \;& I\!P\left[t_1 \leq T_i \leq t_2|X^{i\infty}\right], \quad \forall\, t_1, t_2 = 0, 1, 2...\ (t_1 \leq t_2) \end{aligned} \tag{F-2}$$

et

$$I\!P\left[T_i \geq t|X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] = I\!P\left[T_i \geq t|X^{i\infty}\right], \quad \forall\, t = 0, 1, 2, ... \tag{F-3}$$

- Preuve que (F-1) implique (F-2).

De façon semblable au cas de la preuve faite dans l'Annexe E ci-dessus, $\forall\, t_1, t_2 =$

$0,1,2...$ $(t_1 \leq t_2)$, on a :

$$\begin{aligned} & I\!P\left[t_1 \leq T_i \leq t_2 | X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] \\ = & \sum_{t^*=t_1}^{t_2} I\!P\left[T_i = t^* | X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] \end{aligned}$$

soit, étant donné (F-1),

$$\begin{aligned} & I\!P\left[t_1 \leq T_i \leq t_2 | X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] \\ = & \sum_{t^*=t_1}^{t_2} I\!P\left[T_i = t^* | X^{i\infty}\right] \\ = & I\!P\left[t_1 \leq T_i \leq t_2 | X^{i\infty}\right] \end{aligned}$$

CQFD.

- Preuve que (F-1) implique (F-3).

De façon semblable à ci-dessus, $\forall\, t = 0,1,2,...$, on a :

$$\begin{aligned} & I\!P\left[T_i \geq t | X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] \\ = & 1 - I\!P\left[T_i \leq t-1 | X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] \\ = & 1 - \sum_{t^*=0}^{t-1} I\!P\left[T_i = t^* | X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] \end{aligned}$$

soit, étant donné (F-1),

$$\begin{aligned} & I\!P\left[T_i \geq t | X^{i\infty}, B_{i1}, ..., B_{iH}, C_i\right] \\ = & 1 - \sum_{t^*=0}^{t-1} I\!P\left[T_i = t^* | X^{i\infty}\right] \\ = & 1 - I\!P\left[T_i \leq t-1 | X^{i\infty}\right] \\ = & I\!P\left[T_i \geq t | X^{i\infty}\right] \end{aligned}$$

CQFD.

# Annexe G

- Preuve de l'équivalence des conditions d'exogénéité stricte :

$$D\left(X^{\underline{it+1}} | T_i = t-k, E_i = l, X^{it}\right) = D\left(X^{\underline{it+1}} | X^{it}\right), \quad \begin{array}{l} \forall\, t = 0,1,2,... \\ \forall\, k = 0,1,2,...,t \\ \forall\, l = 1,2,...,L \end{array} \qquad \text{(G-1)}$$

et

$$D\left(X^{\underline{it+1}} | T_i = t, E_i = l, X^{it}\right) = D\left(X^{\underline{it+1}} | X^{it}\right), \quad \forall\, t = 0,1,2,...,\ \forall\, l = 1,2,...,L \qquad \text{(G-2)}$$

On doit montrer que (G-2) est une condition nécessaire et suffisante pour (G-1).

Il suffit d'adapter la preuve de l'Annexe A.

# Annexe H

- Preuve que la condition d'exogénéité stricte :

$$D\left(X^{\underline{it+1}}|T_i = t-k, E_i = l, X^{it}\right) = D\left(X^{\underline{it+1}}|X^{it}\right), \quad \begin{array}{l} \forall\, t = 0,1,2,... \\ \forall\, k = 0,1,2,...,t \\ \forall\, l = 1,2,...,L \end{array} \tag{H-1}$$

implique :

$$\mathbb{P}\left[T_i = t, E_i = l|X^{it}, X^{\underline{it+1}}\right] = \mathbb{P}\left[T_i = t, E_i = l|X^{it}\right], \quad \begin{array}{l} \forall\, t = 0,1,2,... \\ \forall\, l = 1,2,...,L \end{array} \tag{H-2}$$

$$\mathbb{P}\left[T_i = t|X^{it}, X^{\underline{it+1}}\right] = \mathbb{P}\left[T_i = t|X^{it}\right], \quad \forall\, t = 0,1,2,... \tag{H-3}$$

$$\mathbb{P}\left[T_i \geq t|X^{it}, X^{\underline{it+1}}\right] = \mathbb{P}\left[T_i \geq t|X^{it}\right], \quad \forall\, t = 0,1,2,... \tag{H-4}$$

$$\begin{aligned} & \mathbb{P}\left[T_i = t|T_i \geq t, X^{it}, X^{\underline{it+1}}\right] \\ = \; & \mathbb{P}\left[T_i = t|T_i \geq t, X^{it}\right], \quad \forall\, t = 0,1,2,... \end{aligned} \tag{H-5}$$

et

$$\begin{aligned} & \mathbb{P}\left[T_i = t, E_i = l|T_i \geq t, X^{it}, X^{\underline{it+1}}\right] \\ = \; & \mathbb{P}\left[T_i = t, E_i = l|T_i \geq t, X^{it}\right], \quad \begin{array}{l} \forall\, t = 0,1,2,... \\ \forall\, l = 1,2,...,L \end{array} \end{aligned} \tag{H-6}$$

- Preuve que (H-1) implique (H-2).

La condition (H-1) signifie que, conditionnellement à $X^{it}$, $X^{\underline{it+1}}$ et $(T_i = t-k, E_i = l)$ sont indépendants :

$$X^{\underline{it+1}} \perp (T_i = t-k, E_i = l) \mid X^{it}, \quad \begin{array}{l} \forall\, t = 0,1,2,... \\ \forall\, k = 0,1,2,...,t \\ \forall\, l = 1,2,...,L \end{array}$$

Cela implique que, $\forall\, t = 0,1,2,...$, $\forall\, k = 0,1,2,...,t$, $\forall\, l = 1,2,...,L$, on a :

$$\mathbb{P}\left[T_i = t-k, E_i = l|X^{it}, X^{\underline{it+1}}\right] = \mathbb{P}\left[T_i = t-k, E_i = l|X^{it}\right] \tag{H-7}$$

soit, en prenant $k = 0$ :

$$\mathbb{P}\left[T_i = t, E_i = l|X^{it}, X^{\underline{it+1}}\right] = \mathbb{P}\left[T_i = t, E_i = l|X^{it}\right]$$

CQFD.

- Preuve que (H-1) implique (H-3), (H-4) et (H-5).

  La condition (H-1) implique que :

  $D\left(X^{\underline{it+1}}|T_i = t-k, X^{it}\right) = D\left(X^{\underline{it+1}}|X^{it}\right), \ \forall t = 0,1,2,..., \ \forall k = 0,1,2,...,t$

  autrement dit la condition d'exogénéité utilisée dans le modèle de base. Or, nous avons prouvé dans l'annexe B que cette dernière condition implique (H-3), (H-4) et (H-5).

- Preuve que (H-1) implique (H-6).

  $\forall\, t = 0,1,2,...$, on a :

$$I\!P\left[T_i = t, E_i = l|T_i \geq t, X^{it}, X^{\underline{it+1}}\right] = \frac{I\!P\left[T_i = t, E_i = l, T_i \geq t|X^{it}, X^{\underline{it+1}}\right]}{I\!P\left[T_i \geq t|X^{it}, X^{\underline{it+1}}\right]}, \ \forall\, t = 0,1,2,...$$

  soit, comme $(T_i = t)$ implique $(T_i \geq t)$,

$$I\!P\left[T_i = t, E_i = l|T_i \geq t, X^{it}, X^{\underline{it+1}}\right] = \frac{I\!P\left[T_i = t, E_i = l|X^{it}, X^{\underline{it+1}}\right]}{I\!P\left[T_i \geq t|X^{it}, X^{\underline{it+1}}\right]}$$

  soit encore, étant donné (H-2), (H-4) et l'implication évoquée ci-dessus,

$$\begin{aligned} I\!P\left[T_i = t, E_i = l|T_i \geq t, X^{it}, X^{\underline{it+1}}\right] &= \frac{I\!P\left[T_i = t, E_i = l|X^{it}\right]}{I\!P\left[T_i \geq t|X^{it}\right]} \\ &= \frac{I\!P\left[T_i = t, E_i = l, T_i \geq t|X^{it}\right]}{I\!P\left[T_i \geq t|X^{it}\right]} \\ &= I\!P\left[T_i = t, E_i = l|T_i \geq t, X^{it}\right] \end{aligned}$$

  CQFD.

# Annexe I

- Preuve que la condition d'indépendance conditionnelle :

$$I\!P\left[T_i = t, E_i = l|X^{i\infty}, C_i\right] = I\!P\left[T_i = t, E_i = l|X^{i\infty}\right], \quad \begin{array}{l}\forall\, t = 0,1,2,...\\ \forall\, l = 1,2,...,L\end{array} \qquad \text{(I-1)}$$

  implique :

$$I\!P\left[T_i \geq t|X^{i\infty}, C_i\right] = I\!P\left[T_i \geq t|X^{i\infty}\right], \quad \forall\, t = 0,1,2,...$$

  $\forall\, t = 0,1,2,...$, on a :

$$\begin{aligned} I\!P\left[T_i \geq t|X^{i\infty}, C_i\right] &= 1 - I\!P\left[T_i \leq t-1|X^{i\infty}, C_i\right] \\ &= 1 - \sum_{t^*=0}^{t-1} I\!P\left[T_i = t^*|X^{i\infty}, C_i\right] \\ &= 1 - \sum_{t^*=0}^{t-1}\sum_{l=1}^{L} I\!P\left[T_i = t^*, E_i = l|X^{i\infty}, C_i\right] \end{aligned}$$

soit, étant donné (I-1),

$$
\begin{aligned}
I\!\!P\left[T_i \geq t | X^{it}, C_i\right] &= 1 - \sum_{t^*=0}^{t-1} \sum_{l=1}^{L} I\!\!P\left[T_i = t^*, E_i = l | X^{i\infty}\right] \\
&= 1 - \sum_{t^*=0}^{t-1} I\!\!P\left[T_i = t^* | X^{i\infty}\right] \\
&= 1 - I\!\!P\left[T_i \leq t-1 | X^{i\infty}\right] \\
&= I\!\!P\left[T_i \geq t | X^{i\infty}\right]
\end{aligned}
$$

CQFD.