Modèle discret de type 'Timing of Events'

Note de travail

Bernard Lejeune
HEC - Université de Liège

Octobre 2013

On s'intéresse à l'effet d'un traitement – qui peut avoir lieu à différent moment – sur le temps qui s'écoule avant qu'un certain évènement se produise. De façon concrète, on considère le temps qui s'écoule avant qu'un chômeur trouve un emploi, avec comme traitement le fait de suivre un programme de formation.

1. Variables et notations

On note :

- T_u = la durée de chômage (mesurée en mois, semaines ou jours) d'un individu pris au hasard dans une population de gens qui débutent un épisode de chômage. Les valeurs possibles de T_u sont supposées être $t = 0, 1, 2, ...$, $t = 0$ 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...

- T_p = la durée (dans les même unités de mesures) avant une éventuelle entrée dans un programme de formation du même un individu pris au hasard. Les valeurs possibles de T_p sont de façon semblable supposées être $t = 0, 1, 2, ...$ T_p ne peut en pratique être observé que si $T_p \leq T_u$.

- X = un vecteur de variables explicatives (supposées, par commodité, invariantes dans le temps[1]) caractérisant l'individu.

- $V = (V_u, V_p)$ = un vecteur de variables représentant une hétérogénéité non

[1] En pratique, les variables explicatives peuvent être variables dans le temps, pour peu qu'elles soient strictement exogènes.

observée. V est supposé indépendant de X : $V \perp X$.

On cherche à modéliser l'effet causal d'une formation réalisée en $T_p = t_p$ sur la durée de chômage T_u d'un individu pris au hasard, étant donné les caractéristiques X de l'individu, et en prenant en compte une possible hétérogénéité non observée $V = (V_u, V_p)$.

2. Structure du modèle

Un modèle paramétrique est spécifié pour la distribution jointe conditionnelle $T_u, T_p|X, V$ et pour la distribution marginale de l'hétérogénéité non observée V (supposée indépendante de X). De ces deux distributions peut être déduite la distribution jointe non conditionnelle à l'hétérogénéité non observée $T_u, T_p|X$.

La distribution jointe conditionnelle $T_u, T_p|X, V$ décrit la façon dont les durées T_u, T_p dépendent des variables observées X et non observées V. Cette distribution jointe peut être décomposée en une distribution conditionnelle $T_u|T_p, X, V$ et une distribution marginale $T_p|X, V$. La première rend compte de la durée de chômage étant donné la réalisation éventuelle d'une formation, tandis que seconde décrit le processus d'entrée en formation.

La distribution conditionnelle $T_u|T_p, X, V$ et la distribution marginale $T_p|X, V$ sont modélisées au travers de fonctions de hasard conditionnelles $h_u(t|t_p, X, V)$ et $h_p(t|X, V)$.

2.1. Fonctions de hasard conditionnelles $h_u(t|t_p, X, V)$ et $h_p(t|X, V)$

Les fonctions de hasard conditionnelles sont définies par :

$$\begin{aligned} h_u(t|t_p, X, V) &= I\!P\left[T_u = t | T_u \geq t, T_p = t_p, X, V\right], \quad \forall\, t = 0, 1, 2, ... \\ &= \exp\left[\lambda_u(t) + \phi_u(X) + \delta(t|t_p, X) \cdot I(t \geq t_p) + V_u\right] \end{aligned} \tag{1}$$

et

$$\begin{aligned} h_p(t|X, V) &= I\!P\left[T_p = t | T_p \geq t, X, V\right], \quad \forall\, t = 0, 1, 2, ... \\ &= \exp\left[\lambda_p(t) + \phi_p(X) + V_p\right] \end{aligned} \tag{2}$$

La fonction de hasard conditionnelle $h_u(t|t_p, X, V)$ donne la probabilité que la durée T_u de chômage d'un individu pris au hasard soit (exactement) égale à t, sachant qu'il était toujours au chômage en $t - 1$, qu'il a suivi une formation en $T_p = t_p$, et ses caractéristiques observées X et non observées V. Elle est supposée avoir la même forme que dans un modèle MPH (Mixed Proportional Hazard) standard. Le terme $\lambda_u(t)$ est généralement spécifié comme un polynôme ou une fonction constante par intervalle, et le terme $\phi_u(X)$ comme une combinaison linéaire $X'\gamma_u$. Le terme[2]

[2] $I(t \geq t_p)$ dénote une variable binaire qui prend la valeur 1 si $t \geq t_p$, et 0 sinon.

$\delta(t|t_p, X) \cdot I(t \geq t_p)$ capture l'effet d'une formation en $T_p = t_p$ sur le taux de sortie du chômage. On notera que cet effet peut dépendre de t, t_p et X, mais qu'il ne peut y avoir d'effet que pour $t \geq t_p$, i.e., au plus tôt à partir du début de la formation. Cette hypothèse de non anticipation implique en particulier que, pour tout $t < t_p$, on a :

$$\begin{aligned} h_u(t|t_p, X, V) &= \mathbb{P}\left[T_u = t | T_u \geq t, T_p = t_p, X, V\right] \\ &= \mathbb{P}\left[T_u = t | T_u \geq t, T_p \geq t_p, X, V\right] \\ &= \mathbb{P}\left[T_u = t | T_u \geq t, T_p = \infty, X, V\right] \\ &= \exp\left[\lambda_u(t) + \phi_u(X) + V_u\right] \end{aligned} \tag{3}$$

On notera finalement que le hasard $h_u(t|t_p, X, V)$ n'est supposé dépendre que du terme V_u du vecteur d'hétérogénéité non observée V.

La fonction de hasard conditionnelle $h_p(t|X, V)$ donne la probabilité que la durée T_p avant une entrée en formation d'un individu pris au hasard soit (exactement) égale à t, sachant qu'il n'était toujours pas en formation en $t-1$, et ses caractéristiques observées X et non observées V. De façon semblable à $h_u(t|t_p, X, V)$, elle est supposée avoir la même forme que dans un modèle MPH standard, le terme $\lambda_p(t)$ étant généralement spécifié comme un polynôme ou une fonction constante par intervalle, et le terme $\phi_p(X)$ comme une combinaison linéaire $X'\gamma_p$. On notera encore que le hasard $h_p(t|X, V)$ n'est supposé dépendre que du terme V_p du vecteur d'hétérogénéité non observée V.

2.2. Distributions conditionnelles $T_u|T_p, X, V$ et $T_p|X, V$

La fonction de hasard conditionnelle $h_u(t|t_p, X, V)$ caractérise de façon complète la distribution conditionnelle $T_u|T_p, X, V$. Sa fonction de densité est donnée par[3] :

$$\begin{aligned} f_u(t|t_p, X, V) &= \mathbb{P}\left[T_u = t | T_p = t_p, X, V\right], \quad \forall\, t = 0, 1, 2, ... \\ &= \mathbb{P}\left[T_u = t | T_u \geq t, T_p = t_p, X, V\right] \mathbb{P}\left[T_u \geq t | T_p = t_p, X, V\right] \\ &= h_u(t|t_p, X, V) S_u(t|t_p, X, V) \end{aligned} \tag{4}$$

où $S_u(t|t_p, X, V)$ désigne sa fonction de survie, qui est donnée par :

$$\begin{aligned} S_u(t|t_p, X, V) &= \mathbb{P}\left[T_u \geq t | T_p = t_p, X, V\right] \\ &= \begin{cases} 1 & \text{, si } t = 0 \\ \displaystyle\prod_{t^*=0}^{t-1} \left(1 - h_u(t^*|t_p, X, V)\right) & , \forall\, t = 1, 2, ... \end{cases} \end{aligned} \tag{5}$$

L'expression (5) de la fonction de survie $S_u(t|t_p, X, V)$ s'obtient par récurrence.

[3] On note que comme l'événement $(T_u = t)$ implique l'événement $(T_u \geq t)$, et donc que l'intersection de ces deux événements est tout simplement l'événement $(T_u = t)$, on a :

$$\mathbb{P}\left[T_u = t | T_p = t_p, X, V\right] = \mathbb{P}\left[T_u = t, T_u \geq t | T_p = t_p, X, V\right]$$

Pour $t = 0$, on a :

$$S_u(0|t_p, X, V) = I\!P[T_u \geq 0|T_p = t_p, X, V] = 1$$

Pour $t = 1$, on obtient[4] :

$$\begin{aligned}
S_u(1|t_p, X, V) &= I\!P[T_u \geq 1|T_p = t_p, X, V] \\
&= I\!P[T_u \geq 1|T_u \geq 0, T_p = t_p, X, V]\, I\!P[T_u \geq 0|T_p = t_p, X, V] \\
&= (1 - I\!P[T_u < 1|T_u \geq 0, T_p = t_p, X, V])\, I\!P[T_u \geq 0|T_p = t_p, X, V] \\
&= (1 - I\!P[T_u = 0|T_u \geq 0, T_p = t_p, X, V])\, I\!P[T_u \geq 0|T_p = t_p, X, V] \\
&= (1 - h_u(0|t_p, X, V)) \times 1 = 1 - h_u(0|t_p, X, V)
\end{aligned}$$

De même, pour $t = 2$, on a[5] :

$$\begin{aligned}
S_u(2|t_p, X, V) &= I\!P[T_u \geq 2|T_p = t_p, X, V] \\
&= I\!P[T_u \geq 2|T_u \geq 1, T_p = t_p, X, V]\, I\!P[T_u \geq 1|T_p = t_p, X, V] \\
&= (1 - I\!P[T_u < 2|T_u \geq 1, T_p = t_p, X, V])\, I\!P[T_u \geq 1|T_p = t_p, X, V] \\
&= (1 - I\!P[T_u = 1|T_u \geq 1, T_p = t_p, X, V])\, I\!P[T_u \geq 1|T_p = t_p, X, V] \\
&= (1 - h_u(1|t_p, X, V))\,(1 - h_u(0|t_p, X, V))
\end{aligned}$$

et ainsi de suite.

La fonction de densité conditionnelle $f_u(t|t_p, X, V)$ donne la probabilité que la durée T_u de chômage d'un individu pris au hasard soit (exactement) égale à t, sachant qu'il a suivi une formation en $T_p = t_p$, et ses caractéristiques observées X et non observées V. Pour sa part, la fonction de survie conditionnelle $S_u(t|t_p, X, V)$ donne la probabilité que la durée T_u de chômage soit supérieure ou égale à t, sachant qu'il a suivi une formation en $T_p = t_p$, et ses caractéristiques observées X et non observées V. On notera que puisque $h_u(t|t_p, X, V)$ ne dépend en réalité que du terme V_u du vecteur d'hétérogénéité non observée V, il en est de même pour $f_u(t|t_p, X, V)$ et $S_u(t|t_p, X, V)$. On notera finalement que l'hypothèse de non anticipation évoquée précédemment – i.e., il ne peut y avoir d'effet d'une formation en $T_p = t_p$ sur le taux de sortie du chômage que pour $t \geq t_p$ — implique, en conséquence de l'équation (3) que pour tout $t \leq t_p$ (et pas seulement $t < t_p$), on a :

$$\begin{aligned}
S_u(t|t_p, X, V) &= I\!P[T_u \geq t|T_p = t_p, X, V] \\
&= I\!P[T_u \geq t|T_p \geq t_p, X, V] \qquad (6)
\end{aligned}$$

[4] On note que comme l'événement $(T_u \geq 1)$ implique l'événement $(T_u \geq 0)$, et donc que l'intersection de ces deux événements est tout simplement l'événement $(T_u \geq 1)$, on a :

$$I\!P[T_u \geq 1|T_p = t_p, X, V] = I\!P[T_u \geq 1, T_u \geq 0|T_p = t_p, X, V]$$

[5] Pour les même raisons que ci-dessus, on a :

$$I\!P[T_u \geq 2|T_p = t_p, X, V] = I\!P[T_u \geq 2, T_u \geq 1|T_p = t_p, X, V]$$

tandis que pour tout $t < t_p$ (seulement), on a :

$$\begin{aligned} f_u(t|t_p, X, V) &= \mathbb{P}[T_u = t|T_p = t_p, X, V] \\ &= \mathbb{P}[T_u = t|T_p \geq t_p, X, V] \end{aligned} \tag{7}$$

De son côté, la fonction de hasard conditionnelle $h_p(t|X, V)$ caractérise de façon complète la distribution conditionnelle $T_p|X, V$. Sa fonction de densité est donnée par[6] :

$$\begin{aligned} f_p(t|X, V) &= \mathbb{P}[T_p = t|X, V], \quad \forall\, t = 0, 1, 2, ... \\ &= \mathbb{P}[T_p = t|T_p \geq t, X, V]\, \mathbb{P}[T_p \geq t|X, V] \\ &= h_p(t|X, V) S_p(t|X, V) \end{aligned} \tag{8}$$

où $S_p(t|X, V)$ désigne sa fonction de survie, qui est donnée par :

$$\begin{aligned} S_p(t|X, V) &= \mathbb{P}[T_p \geq t|X, V] \\ &= \begin{cases} 1 & , \text{si } t = 0 \\ \prod_{t^*=0}^{t-1} (1 - h_p(t^*|X, V)) & , \forall\, t = 1, 2, ... \end{cases} \end{aligned} \tag{9}$$

L'expression (9) de la fonction de survie $S_p(t|X, V)$ s'obtient de façon semblable à $S_u(t|t_p, X, V)$ par récurrence.

La fonction de densité conditionnelle $f_p(t|X, V)$ donne la probabilité que la durée T_p avant une entrée en formation d'un individu pris au hasard soit (exactement) égale à t, sachant ses caractéristiques observées X et non observées V. Pour sa part, la fonction de survie conditionnelle $S_p(t|X, V)$ donne la probabilité que la durée T_p avant une entrée en formation soit supérieure ou égale à t, sachant ses caractéristiques observées X et non observées V. On notera encore que puisque $h_p(t|X, V)$ ne dépend en réalité que du terme V_p du vecteur d'hétérogénéité non observée V, il en est de même pour $f_p(t|X, V)$ et $S_p(t|X, V)$.

2.3. Distribution jointe conditionnelle $T_u, T_p|X, V$

Les distributions conditionnelles $T_u|T_p, X, V$ et $T_p|X, V$ caractérisent à leur tour de façon complète la distribution jointe conditionnelle $T_u, T_p|X, V$. Sa fonction de densité est donnée par :

$$\begin{aligned} f(t_u, t_p|X, V) &= \mathbb{P}[T_u = t_u, T_p = t_p|X, V], \quad \forall\, t_u, t_p = 0, 1, 2, ... \\ &= \mathbb{P}[T_u = t_u|T_p = t_p, X, V]\, \mathbb{P}[T_p = t_p|X, V] \\ &= f_u(t_u|t_p, X, V) f_p(t_p|X, V) \end{aligned} \tag{10}$$

---

[6] On note que, pour la même raison que celle précédemment évoquée, on a :

$$\mathbb{P}[T_p = t|X, V] = \mathbb{P}[T_p = t, T_p \geq t|X, V]$$

La fonction de densité jointe conditionnelle $f(t_u, t_p|X, V)$ donne la probabilité que la durée T_u de chômage et la durée T_p avant une entrée en formation d'un individu pris au hasard soient (exactement) égales à respectivement t_u et t_p, sachant ses caractéristiques observées X et non observées V. On notera que sachant X, V, les durées T_u et T_p ne sont dépendantes qu'au travers de l'effet d'une réalisation de T_p en t_p sur le taux de sortie du chômage, i.e., par l'intermédiare du terme $\delta(t|t_p, X) \cdot I(t \geq t_p)$ qui apparaît dans la fonction de hasard conditionnelle $h_u(t|t_p, X, V)$.

2.3.1. Distribution de l'hétérogénéité non observée $V = (V_u, V_p)$

On spécifie généralement une distribution discrète pour la distribution marginale de $V = (V_u, V_p)$, qui par ailleurs est supposée indépendante de X. Sa fonction de densité est par définition donnée par :

$$\begin{aligned} g(v) &= g(v_u, v_p) = I\!P[V_u = v_u, V_p = v_p] \\ &= I\!P[V = v] \end{aligned} \tag{11}$$

L'indépendance de V par rapport à X implique que :

$$I\!P[V = v|X] = I\!P[V = v] = g(v)$$

La fonction de densité $g(v)$ donne la probabilité que la valeur $V = (V_u, V_p)$ de l'hétérogénéité non observée d'un individu pris au hasard soit égale à $v = (v_u, v_p)$. Elle indique la distribution de l'hétérogénéité non observée dans la population des gens qui débutent un épisode de chômage, i.e., en $t = 0$, dans le flux d'entrée au chômage.

2.4. Distribution jointe non conditionnelle à l'hétérogénéité non observée $T_u, T_p|X$

La distribution jointe conditionnelle $T_u, T_p|X, V$ et la distribution distribution marginale de $V = (V_u, V_p)$ caractérise finalement la distribution jointe non conditionnelle à l'hétérogénéité non observée $T_u, T_p|X$. Sa fonction de densité est donnée par :

$$\begin{aligned} f(t_u, t_p|X) &= I\!P[T_u = t_u, T_p = t_p|X], \quad \forall\, t_u, t_p = 0, 1, 2, ... \\ &= \sum_{\{v\}} I\!P[T_u = t_u, T_p = t_p|X, V = v]\, I\!P[V = v|X] \\ &= \sum_{\{v\}} I\!P[T_u = t_u, T_p = t_p|X, V = v]\, I\!P[V = v] \\ &= \sum_{\{v\}} f(t_u, t_p|X, v) g(v) = \sum_{\{v\}} f_u(t_u|t_p, X, v) f_p(t_p|X, v) g(v) \\ &= \sum_{\{v\}} h_u(t_u|t_p, X, v) S_u(t_u|t_p, X, v) h_p(t_p|X, v) S_p(t_p|X, v) g(v) \end{aligned} \tag{12}$$

La fonction de densité jointe conditionnelle $f(t_u, t_p|X)$ donne la probabilité que la durée T_u de chômage et la durée T_p avant une entrée en formation d'un individu pris au hasard soient (exactement) égales à respectivement t_u et t_p, sachant ses caractéristiques observées X, mais non conditionnellement à l'hétérogénéité non observées V. On notera que conditionnellement à X seulement, la dépendance entre les durées T_u et T_p provient à la fois de l'effet d'une réalisation de T_p en t_p sur le taux de sortie du chômage – par l'intermédiare du terme $\delta(t|t_p, X) \cdot I(t \geq t_p)$ qui apparaît dans (1) – et d'une possible dépendance entre les termes V_u et V_p du vecteur d'hétérogénéité non observée V.

Les valeurs de T_u et de T_p ne sont pas toujours observées de façon exacte. En pratique, T_p ne peut être observé de façon exacte que si $T_p \leq T_u$, i.e., si la période d'entrée en formation est inférieure, ou au plus égale, à la période du sortie du chômage. Lorsque ce n'est pas le cas, il est simplement observé que $T_p \geq T_u + 1$. La probabilité conditionnelle à X, mais non conditionnelles à V, que la durée T_u de chômage soit (exactement) égale à t_u et que $T_p \geq t_u + 1$ est donnée par :

$$\begin{aligned}
& I\!P[T_u = t_u, T_p \geq t_u + 1|X], \quad \forall\, t_u = 0, 1, 2, ... \\
= & \sum_{\{v\}} I\!P[T_u = t_u, T_p \geq t_u + 1|X, V = v]\, I\!P[V = v|X] \\
= & \sum_{\{v\}} I\!P[T_u = t_u|T_p \geq t_u + 1, X, V = v]\, I\!P[T_p \geq t_u + 1|X, V = v]\, I\!P[V = v]
\end{aligned}$$

On sait que :

$$I\!P[T_p \geq t_u + 1|X, V = v] = S_p(t_u + 1|X, v)$$

Comme par définition $t_u < t_u + 1$, de l'équation (7), on a :

$$\begin{aligned}
& I\!P[T_u = t_u|T_p \geq t_u + 1, X, V = v] \\
= & I\!P[T_u = t_u|T_p = t_u + 1, X, V = v] \\
= & f_u(t_u|t_u + 1, X, v) = h_u(t_u|t_u + 1, X, v) S_u(t_u|t_u + 1, X, v)
\end{aligned}$$

de sorte qu'au total on a :

$$\begin{aligned}
& I\!P[T_u = t_u, T_p \geq t_u + 1|X], \quad \forall\, t_u = 0, 1, 2, ... \\
= & \sum_{\{v\}} h_u(t_u|t_u + 1, X, v) S_u(t_u|t_u + 1, X, v) S_p(t_u + 1|X, v) g(v) \qquad (13)
\end{aligned}$$

Généralement, un épisode de chômage n'est observé que sur un nombre fini de périodes. Soit C la dernière période d'observation (= la période de censure à droite). Dans cette situation, la valeur de T_u ne peut être observée de façon exacte que si $T_u \leq C$. Lorsque ce n'est pas le cas, il est simplement observé que $T_u \geq C + 1$, et la valeur de T_p peut ou non être observée de façon exacte selon que $T_p \leq C$ ou non. Si ce n'est pas le cas, il est de manière semblable simplement observé que $T_p \geq C + 1$. La probabilité conditionnelle à X, mais non conditionnelles à V, que $T_u \geq c + 1$ et que $T_p = t_p$ est donnée par :

$$\begin{aligned} & I\!\!P[T_u \geq c+1, T_p = t_p | X], \quad \forall\, c, t_p = 0,1,2,... \\ = & \sum_{\{v\}} I\!\!P[T_u \geq c+1, T_p = t_p | X, V = v]\, I\!\!P[V = v | X] \\ = & \sum_{\{v\}} I\!\!P[T_u \geq c+1 | T_p = t_p, X, V = v]\, I\!\!P[T_p = t_p | X, V = v]\, I\!\!P[V = v] \end{aligned}$$

On sait que :

$$I\!\!P[T_u \geq c+1 | T_p = t_p, X, V = v] = S_u(c+1 | t_p, X, v)$$

$$I\!\!P[T_p = t_p | X, V = v] = f_p(t_p | X, v) = h_p(t_p | X, v) S_p(t_p | X, v)$$

de sorte qu'au total on a :

$$\begin{aligned} & I\!\!P[T_u \geq c+1, T_p = t_p | X], \quad \forall\, c, t_p = 0,1,2,... \\ = & \sum_{\{v\}} S_u(c+1 | t_p, X, v) h_p(t_p | X, v) S_p(t_p | X, v) g(v) \end{aligned} \tag{14}$$

De son côté, la probabilité conditionnelle à X, mais non conditionnelles à V, que $T_u \geq c+1$ et que $T_p \geq c+1$ est donnée par :

$$\begin{aligned} & I\!\!P[T_u \geq c+1, T_p \geq c+1 | X], \quad \forall\, c = 0,1,2,... \\ = & \sum_{\{v\}} I\!\!P[T_u \geq c+1, T_p \geq c+1 | X, V = v]\, I\!\!P[V = v | X] \\ = & \sum_{\{v\}} I\!\!P[T_u \geq c+1 | T_p \geq c+1, X, V = v]\, I\!\!P[T_p \geq c+1 | X, V = v]\, I\!\!P[V = v] \end{aligned}$$

On sait que :

$$I\!\!P[T_p \geq c+1 | X, V = v] = S_p(c+1 | X, v)$$

Comme par définition $c+1 \leq c+1$, de l'équation (6), on a :

$$\begin{aligned} & I\!\!P[T_u \geq c+1 | T_p \geq c+1, X, V = v] \\ = & I\!\!P[T_u \geq c+1 | T_p = c+1, X, V = v] \\ = & S_u(c+1 | c+1, X, v) \end{aligned}$$

de sorte qu'au total on a :

$$\begin{aligned} & I\!\!P[T_u \geq c+1, T_p \geq c+1 | X], \quad \forall\, c = 0,1,2,... \\ = & \sum_{\{v\}} S_u(c+1 | c+1, X, v) S_p(c+1 | X, v) g(v) \end{aligned} \tag{15}$$

2.5. Distribution conditionnelle de l'hétérogénéité non observée

La fonction de densité $g(v) = g(v_u, v_p)$ indique la distribution de l'hétérogénéité non observée $V = (V_u, V_p)$ dans l'ensemble de la population des gens qui débutent un épisode de chômage, i.e., pour un individu pris au hasard en $t = 0$, dans le flux d'entrée au chômage. Suite à un processus d'écrémage (ceux qui ont des taux de sortie plus élévés sortent plus rapidement), si on considère un individu pris au hasard parmi une sous-population de ces individus qui ont une certaine durée de chômage, qui ont suivi une formation à une certaine période et qui ont des caractéristiques observables X particulières, la distribution de V n'est plus la même. Dans la perspective du calcul de l'effet agrégé du traitement pour les individus traités (i.e., qui ont suivis une formation), il est utile de dériver la distribution de l'hétérogénéité non observée V conditionnellement au fait (1) d'avoir suivi une formation en $T_p = t_p$, (2) d'avoir une durée de chômage $T_u \geq t_p$ (i.e., de ne pas être sorti du chômage avant la formation) et (3) d'avoir des caractéristiques observées X. Cette distribution conditionnelle $V|T_u \geq t_p, T_p = t_p, X$ s'obtient par application du théorème de Bayes :

$$\begin{aligned} g^R(v|t_p, X) &= \mathbb{P}[V = v|T_u \geq t_p, T_p = t_p, X] \\ &= \frac{\mathbb{P}[T_u \geq t_p, T_p = t_p|X, V = v]\ \mathbb{P}[V = v|X]}{\mathbb{P}[T_u \geq t_p, T_p = t_p|X]} \end{aligned}$$

On sait que :

$$\mathbb{P}[V = v|X] = \mathbb{P}[V = v] = g(v)$$

On a par ailleurs :

$$\begin{aligned} & \mathbb{P}[T_u \geq t_p, T_p = t_p|X, V = v] \\ = \; & \mathbb{P}[T_u \geq t_p|T_p = t_p, X, V = v]\ \mathbb{P}[T_p = t_p|X, V = v] \\ = \; & S_u(t_p|t_p, X, v)h_p(t_p|X, v)S_p(t_p|X, v) \end{aligned}$$

et

$$\begin{aligned} & \mathbb{P}[T_u \geq t_p, T_p = t_p|X] \\ = \; & \sum_{\{v^*\}} \mathbb{P}[T_u \geq t_p, T_p = t_p|X, V = v^*]\ \mathbb{P}[V = v^*|X] \\ = \; & \sum_{\{v^*\}} \mathbb{P}[T_u \geq t_p, T_p = t_p|X, V = v^*]\ \mathbb{P}[V = v^*] \end{aligned}$$

de sorte qu'au total on a:

$$\begin{aligned} g^R(v|t_p, X) &= \mathbb{P}[V = v|T_u \geq t_p, T_p = t_p, X] \\ &= \frac{S_u(t_p|t_p, X, v)h_p(t_p|X, v)S_p(t_p|X, v)g(v)}{\sum_{\{v^*\}} S_u(t_p|t_p, X, v^*)h_p(t_p|X, v^*)S_p(t_p|X, v^*)g(v^*)} \end{aligned} \tag{16}$$

3. Estimation du modèle

On cherche à estimer les fonctions de hasard conditionnelles $h_u(t|t_p, X, V)$ et $h_p(t|X, V)$ d'un individu pris au hasard parmi les individus d'une population cible (les individus qui s'inscrivent comme demandeur d'emploi au cours d'une période donnée).

On commence par spécifier des formes fonctionnelles $h_u(t|t_p, X, V; \beta_u)$ et $h_p(t|X, V; \beta_p)$, qui dépendent de vecteurs de paramètres $\beta_u \in \Theta_u$ et $\beta_p \in \Theta_p$, pour les fonctions de hasard conditionnelles $h_u(t|t_p, X, V)$ et $h_p(t|X, V)$, formes fonctionnelles que l'on suppose correctement spécifiées, i.e. telle que :

$$\begin{aligned} h_u(t|t_p, X, V; \beta_u^o) &= h_u(t|t_p, X, V), \text{ pour une valeur } \beta_u^o \in \Theta_u \\ h_p(t|X, V; \beta_p^o) &= h_p(t|X, V), \text{ pour une valeur } \beta_p^o \in \Theta_p \end{aligned}$$

Aux fonctions de hasard conditionnelles $h_u(t|t_p, X, V; \beta_u)$ et $h_p(t|X, V; \beta_p)$ correspondent les fonctions de survies :

$$S_u(t|t_p, X, V; \beta_u) = \begin{cases} 1 & \text{, si } t = 0 \\ \prod_{t^*=0}^{t-1} (1 - h_u(t^*|t_p, X, V; \beta_u)) & \text{, } \forall\, t = 1, 2, ... \end{cases}$$

et

$$S_p(t|X, V; \beta_p) = \begin{cases} 1 & \text{, si } t = 0 \\ \prod_{t^*=0}^{t-1} \left(1 - h_p(t^*|X, V; \beta_p)\right) & \text{, } \forall\, t = 1, 2, ... \end{cases}$$

qui, si les fonctions de hasard sont correctement spécifiées, soit elles-mêmes correctement spécifiées, i.e. telle que :

$$\begin{aligned} S_u(t|t_p, X, V; \beta_u^o) &= S_u(t|t_p, X, V), \text{ pour } \beta_u^o \in \Theta_u \\ S_p(t|X, V; \beta_p^o) &= S_p(t|X, V), \text{ pour } \beta_p^o \in \Theta_p \end{aligned}$$

De même, une forme fonctionnelle $g(v; \alpha)$, qui dépend d'un vecteur de paramètres $\alpha \in \Theta_\alpha$, est spécifiée pour la distribution $g(v)$ de l'hétérogénéité non observée, forme fonctionnelle que l'on suppose à nouveau correctement spécifiée, i.e. telle que :

$$g(v; \alpha^o) = g(v), \text{ pour une valeur } \alpha^o \in \Theta_\alpha$$

Une distribution bivariée discrète, entièrement non contrainte, est généralement spécifiée pour $g(v; \alpha)$. Dans ce cas, le vecteur de paramètres α contient, d'une part, les différentes valeurs possibles de v_u et de v_p, et d'autre part, les probabilités des couples de valeurs possibles $v = (v_u, v_p)$.

3.1. Echantillonnage

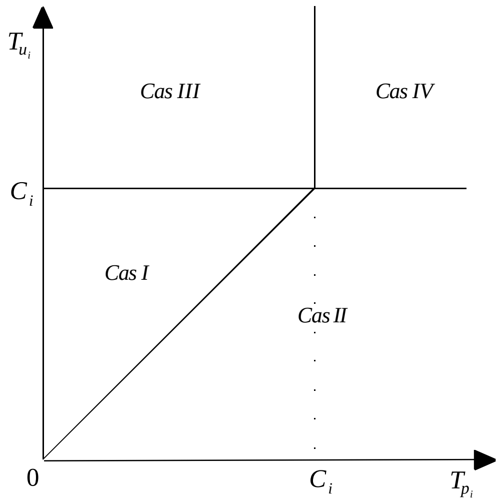
On suppose que l'estimation s'appuye sur un échantillon d'observations de taille n obtenu en tirant au hasard n individus dans la population cible. On note T_{u_i}, T_{p_i}

et X_i respectivement la durée de chômage, la durée avant entrée en formation et les caractéristiques observées d'un individu i pris au hasard. On suppose par ailleurs que l'épisode de chômage de chaque individu i est observé durant une période allant de $t = 0$ jusque $t = C_i$, où C_i désigne un point de censure à droite[7].

Etant donné la fenêtre d'observation finie (censure à droite) et le fait que T_{p_i} ne peut être observé que si $T_{p_i} \leq T_{u_i}$, les durées T_{u_i} et T_{p_i} ne sont pas toujours observées de façon exactes. On peut distinguer 4 cas :

- Cas I : $T_{u_i} \leq C_i$ et $T_{p_i} \leq T_{u_i}$ Notons que ces conditions impliquent qu'on a aussi $T_{p_i} \leq C_i$. Dans ce cas, tant T_{u_i} que T_{p_i} sont observés de façon exacte.
- Cas II : $T_{u_i} \leq C_i$ et $T_{p_i} > T_{u_i}$. Dans ce cas, T_{u_i} est observé de façon exacte, mais T_{p_i} est censuré : il est simplement observé que $T_{p_i} \geq T_{u_i} + 1$.
- Cas III : $T_{u_i} > C_i$ et $T_{p_i} \leq C_i$. On note que ces conditions impliquent qu'on a aussi $T_{p_i} \leq T_{u_i}$. Dans ce cas, T_{p_i} est observé de façon exacte, mais T_{u_i} est censuré : il est simplement observé que $T_{u_i} \geq C_i + 1$.
- Cas IV : $T_{u_i} > C_i$ et $T_{p_i} > C_i$. Dans ce cas, tant T_{u_i} que T_{p_i} sont censurés : il est simplement observé que $T_{u_i} \geq C_i + 1$ et que $T_{p_i} \geq C_i + 1$.

Graphiquement :



Graphique 1 : schéma d'échantillonnage

Pour ce schéma d'échantillonnage, on peut représenter ce qui est en pratique observé comme une réalisation du quintuple $(T^*_{u_i}, T^*_{p_i}, D_{u_i}, D_{p_i}, X_i)$, où :

- D_{u_i} est une variable binaire (auxiliaire) qui prend la valeur 1 si T_{u_i} est observé de façon exacte, et 0 sinon (T_{u_i} est censuré).
- D_{p_i} est une variable binaire (auxiliaire) qui prend la valeur 1 si T_{p_i} est observé de façon exacte, et 0 sinon (T_{p_i} est censuré).
- $T^*_{u_i}$ et $T^*_{p_i}$ désignent respectivement la durée de chômage et la durée avant entrée en formation "effectivement" observée de l'individu i étant donné le mécanisme

[7] Dans la suite, il est implicitement supposé que le mécanisme de censure est 'ignorable' (ou 'exogène'), i.e., que $(T_{u_i}, T_{p_i}) \perp C_i \,|X_i, V_i$ et que $V_i \perp (X_i, C_i)$.

d'échantillonnage, durées qui sont égale à :

$$(T^*_{u_i}, T^*_{p_i}) = \begin{cases} (T_{u_i}, T_{p_i}) & \text{, si } D_{u_i} = 1 \text{ et } D_{p_i} = 1 \text{ (= Cas I)} \\ (T_{u_i}, T_{u_i} + 1) & \text{, si } D_{u_i} = 1 \text{ et } D_{p_i} = 0 \text{ (= Cas II)} \\ (C_i + 1, T_{p_i}) & \text{, si } D_{u_i} = 0 \text{ et } D_{p_i} = 1 \text{ (= Cas III)} \\ (C_i + 1, C_i + 1) & \text{, si } D_{u_i} = 0 \text{ et } D_{p_i} = 0 \text{ (= Cas IV)} \end{cases}$$

où $T^*_{u_i}$ et $T^*_{p_i}$ sont par convention fixés à la valeur minimale (= période de censure $+1$, i.e., selon les cas, $T_{u_i} + 1$ ou $C_i + 1$) des durées T_{u_i} et T_{p_i} lorsque celles-ci sont censurées.

3.2. Estimateur du maximum de vraisemblance

L'estimateur maximum de vraisemblance du modèle est défini par :

$$\left(\hat{\beta}_u, \hat{\beta}_p, \hat{\alpha}\right) = \text{Argmax}_{(\beta_u, \beta_p, \alpha)} \sum_{i=1}^{n} \ln f^*(T^*_{u_i}, T^*_{p_i}, D_{u_i}, D_{p_i} | X_i; \beta_u, \beta_p, \alpha) \tag{17}$$

où :

$$\begin{aligned} & f^*(t^*_{u_i}, t^*_{p_i}, d_{u_i}, d_{p_i} | x_i; \beta_u, \beta_p, \alpha) \\ = \; & I\!P\left[T^*_{u_i} = t^*_{u_i}, T^*_{p_i} = t^*_{p_i}, D_{u_i} = d_{u_i}, D_{p_i} = d_{p_i} | X_i = x_i; \beta_u, \beta_p, \alpha\right] \end{aligned} \tag{18}$$

est la fonction de densité jointe, non conditionnelle à l'hétérogénéité non observée, de $T^*_{u_i}, T^*_{p_i}, D_{u_i}, D_{p_i} | X_i$.

Lorsque $d_{u_i} = d_{p_i} = 1$ (= Cas I), on a :

$$\begin{aligned} & I\!P\left[T^*_{u_i} = t^*_{u_i}, T^*_{p_i} = t^*_{p_i}, D_{u_i} = 1, D_{p_i} = 1 | X_i = x_i; \beta_u, \beta_p, \alpha\right] \\ = \; & I\!P\left[T_{u_i} = t^*_{u_i}, T_{p_i} = t^*_{p_i} | X_i = x_i; \beta_u, \beta_p, \alpha\right] \\ = \; & \sum_{\{v\}} h_u(t^*_{u_i} | t^*_{p_i}, x_i, v; \beta_u) S_u(t^*_{u_i} | t^*_{p_i}, x_i, v; \beta_u) \\ & \quad \times h_p(t^*_{p_i} | x_i, v; \beta_p) S_p(t^*_{p_i} | x_i, v; \beta_p) g(v; \alpha) \end{aligned} \tag{19}$$

où la dernière égalité découle de (12).

Lorsque $d_{u_i} = 1$ et $d_{p_i} = 0$ (= Cas II), par définition, $t^*_{p_i} = t^*_{u_i} + 1$, et on a :

$$\begin{aligned} & I\!P\left[T^*_{u_i} = t^*_{u_i}, T^*_{p_i} = t^*_{p_i}, D_{u_i} = 1, D_{p_i} = 0 | X_i = x_i; \beta_u, \beta_p, \alpha\right] \\ = \; & I\!P\left[T_{u_i} = t^*_{u_i}, T_{p_i} \geq t^*_{u_i} + 1 | X_i = x_i; \beta_u, \beta_p, \alpha\right] \\ = \; & \sum_{\{v\}} h_u(t^*_{u_i} | t^*_{u_i} + 1, x_i, v; \beta_u) S_u(t^*_{u_i} | t^*_{u_i} + 1, x_i, v; \beta_u) S_p(t^*_{u_i} + 1 | x_i, v; \beta_p) g(v; \alpha) \\ = \; & \sum_{\{v\}} h_u(t^*_{u_i} | t^*_{p_i}, x_i, v; \beta_u) S_u(t^*_{u_i} | t^*_{p_i}, x_i, v; \beta_u) S_p(t^*_{p_i} | x_i, v; \beta_p) g(v; \alpha) \end{aligned} \tag{20}$$

où l'avant-dernière égalité découle de (13), et la dernière égalité du fait que, par définition, $t^*_{p_i} = t^*_{u_i} + 1$.

Lorsque $d_{u_i} = 0$ et $d_{p_i} = 1$ (= Cas III), par définition, $t^*_{u_i} = c_i + 1$, et on a :

$$\begin{aligned}
& I\!P\left[T^*_{u_i} = t^*_{u_i}, T^*_{p_i} = t^*_{p_i}, D_{u_i} = 0, D_{p_i} = 1 | X_i = x_i; \beta_u, \beta_p, \alpha\right] \\
= \; & I\!P\left[T_{u_i} \geq c_i + 1, T_{p_i} = t^*_{p_i} | X_i = x_i; \beta_u, \beta_p, \alpha\right] \\
= \; & \sum_{\{v\}} S_u(c_i + 1 | t^*_{p_i}, x_i, v; \beta_u) h_p(t^*_{p_i} | x_i, v; \beta_p) S_p(t^*_{p_i} | x_i, v; \beta_p) g(v; \alpha) \\
= \; & \sum_{\{v\}} S_u(t^*_{u_i} | t^*_{p_i}, x_i, v; \beta_u) h_p(t^*_{p_i} | x_i, v; \beta_p) S_p(t^*_{p_i} | x_i, v; \beta_p) g(v; \alpha) \qquad (21)
\end{aligned}$$

où l'avant-dernière égalité découle de (14), et la dernière égalité du fait que, par définition, $t^*_{u_i} = c_i + 1$.

Finalement, lorsque $d_{u_i} = d_{p_i} = 0$ (= Cas IV), par définition, $t^*_{u_i} = t^*_{p_i} = c_i + 1$, et on a :

$$\begin{aligned}
& I\!P\left[T^*_{u_i} = t^*_{u_i}, T^*_{p_i} = t^*_{p_i}, D_{u_i} = 0, D_{p_i} = 0 | X_i = x_i; \beta_u, \beta_p, \alpha\right] \\
= \; & I\!P\left[T_{u_i} \geq c_i + 1, T_{p_i} \geq c_i + 1 | X_i = x_i; \beta_u, \beta_p, \alpha\right] \\
= \; & \sum_{\{v\}} S_u(c_i + 1 | t^*_{p_i}, x_i, v; \beta_u) S_p(c_i + 1 | x_i, v; \beta_p) g(v; \alpha) \\
= \; & \sum_{\{v\}} S_u(t^*_{u_i} | t^*_{p_i}, x_i, v; \beta_u) S_p(t^*_{p_i} | x_i, v; \beta_p) g(v; \alpha) \qquad (22)
\end{aligned}$$

où l'avant-dernière égalité découle de (15), et la dernière égalité du fait que, par définition, $t^*_{u_i} = t^*_{p_i} = c_i + 1$.

Etant donné les expressions (19) - (22), l'estimateur maximum de vraisemblance du modèle peut simplement s'écrire :

$$\left(\hat{\beta}_u, \hat{\beta}_p, \hat{\alpha}\right) = \text{Argmax}_{(\beta_u, \beta_p, \alpha)} \sum_{i=1}^{n} \ln f^*(T^*_{u_i}, T^*_{p_i}, D_{u_i}, D_{p_i} | X_i; \beta_u, \beta_p, \alpha) \qquad (23)$$

où :

$$\begin{aligned}
& f^*(T^*_{u_i}, T^*_{p_i}, D_{u_i}, D_{p_i} | X_i; \beta_u, \beta_p, \alpha) \\
= \; & \sum_{\{v\}} \; \left[h_u(T^*_{u_i} | T^*_{p_i}, X_i, v; \beta_u)\right]^{D_{u_i}} S_u(T^*_{u_i} | T^*_{p_i}, X_i, v; \beta_u) \qquad (24) \\
& \quad \times \; \left[h_p(T^*_{p_i} | X_i, v; \beta_p)\right]^{D_{p_i}} S_p(T^*_{p_i} | X_i, v; \beta_p) g(v; \alpha)
\end{aligned}$$

Sous des conditions de régularité générales, l'estimateur $\hat{\theta} = \left(\hat{\beta}'_u, \hat{\beta}'_p, \hat{\alpha}'\right)'$ est un estimateur convergent et asymptotiquement normal de $\theta^o = \left(\beta^{o\prime}_u, \beta^{o\prime}_p, \alpha^{o\prime}\right)'$:

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

où :

$$\begin{aligned}\Phi_o &= \frac{1}{n}\sum_{i=1}^{n} E\left[\left(\frac{\partial \ln f^*(T^*_{u_i},T^*_{p_i},D_{u_i},D_{p_i}|X_i;\theta^o)}{\partial\theta}\right)\left(\frac{\partial \ln f^*(T^*_{u_i},T^*_{p_i},D_{u_i},D_{p_i}|X_i;\theta^o)}{\partial\theta}\right)'\right] \\ &= -\frac{1}{n}\sum_{i=1}^{n} E\left[\frac{\partial^2 \ln f^*(T^*_{u_i},T^*_{p_i},D_{u_i},D_{p_i}|X_i;\theta^o)}{\partial\theta\partial\theta'}\right]\end{aligned}$$

soit en termes d'approximation utilisable est échantillon fini :

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

où $\hat\Phi$ est un estimateur convergent de Φ_o. On utilise typiquement :

$$\hat\Phi = -\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2 \ln f^*(T^*_{u_i},T^*_{p_i},D_{u_i},D_{p_i}|X_i;\hat\theta)}{\partial\theta\partial\theta'}$$

4. Effet agrégé du traitement pour les individus traités

On cherche à estimer l'effet agrégé du traitement pour les individus traités, en contrastant, d'une part, la durée 'résiduelle' (i.e., après formation) de chômage des individus qui ont suivi une formation, et d'autre part, la durée 'résiduelle' de chômage des mêmes individus qui aurait prévalu s'ils n'avaient pas suivi une formation (= le contrefactuel).

On considère un individu pris au hasard parmi les individus qui ont suivi une formation en $T_p = t_p$, étaient toujours au chômage au début de cette formation ($T_u \geq t_p$), et qui ont des caractéristiques observées X et non observées V. On note $T_u^R = T_u - t_p$ la durée 'résiduelle' de chômage à partir du début de la formation ($T_u^R = 0$ indique une sortie du chômage à la même période que l'entrée en formation, $T_u^R = 1$, une sortie du chômage la première période suivant celle d'entrée en formation, etc...). On a[8,9] :

[8] On note que, $\forall t \geq 0$, l'événement ($T_u = t + t_p$) implique l'événement ($T_u \geq t_p$), et donc que l'intersection de ces deux événements est tout simplement l'événement ($T_u = t + t_p$).

[9] Exprimé en termes de durées potentielles $\{T_u^*(t_p), t_p = 0, 1, ...\}$, on a :

$$f_u^R(t|t_p, X, V) = \mathbb{P}[T_u^*(t_p) = t + t_p | T_u^*(t_p) \geq t_p, T_p = t_p, X, V]\,, \quad \forall t = 0, 1, 2, ...$$

où $T_u^*(t_p)$ dénote la durée de chômage potentielle pour un traitement en t_p, la durée effectivement observée étant $T_u = T_u^*(T_p)$, i.e, la durée de chômage pour le traitement qui est effectivement reçu (un seul des traitements potentiels est effectivement reçu). L'égalité de cette expression avec l'expression en terme de durée effectivement observée ci-dessus découle du fait que lorsque $T_p = t_p$, $T_u = T_u^*(t_p)$.

$$\begin{aligned} f_u^R(t|t_p, X, V) &= I\!P\left[T_u^R = t | T_u \geq t_p, T_p = t_p, X, V\right], \quad \forall t = 0, 1, 2, ... \\ &= I\!P[T_u = t + t_p | T_u \geq t_p, T_p = t_p, X, V] \\ &= \frac{I\!P[T_u = t + t_p, T_u \geq t_p | T_p = t_p, X, V]}{I\!P[T_u \geq t_p | T_p = t_p, X, V]} \\ &= \frac{I\!P[T_u = t + t_p | T_p = t_p, X, V]}{I\!P[T_u \geq t_p | T_p = t_p, X, V]} \end{aligned}$$

De (4) et (5), on sait que :

$$I\!P[T_u = t + t_p | T_p = t_p, X, V] = h_u(t + t_p | t_p, X, V) S_u(t + t_p | t_p, X, V)$$

et

$$I\!P[T_u \geq t_p | T_p = t_p, X, V] = S_u(t_p | t_p, X, V)$$

de sorte qu'au total on a :

$$\begin{aligned} f_u^R(t|t_p, X, V) &= I\!P\left[T_u^R = t | T_u \geq t_p, T_p = t_p, X, V\right], \quad \forall t = 0, 1, 2, ... \\ &= \frac{h_u(t + t_p | t_p, X, V) S_u(t + t_p | t_p, X, V)}{S_u(t_p | t_p, X, V)} \end{aligned} \tag{25}$$

La fonction de densité conditionnelle $f_u^R(t|t_p, X, V)$ donne la probabilité que la durée 'résiduelle' T_u^R de chômage d'un individu pris au hasard soient (exactement) égale t, sachant qu'il a suivi une formation en $T_p = t_p$, qu'il était toujours au chômage au début de cette formation ($T_u \geq t_p$), et ses caractéristiques observées X et non observées V.

Le contrefactuel de cette distribution de probabilité est donné par la fonction de densité conditionnelle[10] :

$$\begin{aligned} \bar{f}_u^R(t|t_p, X, V) &= I\!P\left[T_u^R = t | T_u \geq t_p, T_p = \infty, X, V\right], \quad \forall t = 0, 1, 2, ... \\ &= I\!P[T_u = t + t_p | T_u \geq t_p, T_p = \infty, X, V] \end{aligned}$$

où ($T_p = \infty$) correspond à une absence de formation. On a[11] :

$$\begin{aligned} & I\!P[T_u = t + t_p | T_u \geq t_p, T_p = \infty, X, V] \\ &= \frac{I\!P[T_u = t + t_p, T_u \geq t_p | T_p = \infty, X, V]}{I\!P[T_u \geq t_p | T_p = \infty, X, V]} \\ &= \frac{I\!P[T_u = t + t_p | T_p = \infty, X, V]}{I\!P[T_u \geq t_p | T_p = \infty, X, V]} \end{aligned}$$

[10] Exprimé en termes de durées potentielles $\{T_u^*(t_p), t_p = 0, 1, ...\}$, on a :

$$\bar{f}_u^R(t|t_p, X, V) = I\!P[T_u^*(\infty) = t + t_p | T_u^*(\infty) \geq t_p, T_p = t_p, X, V], \quad \forall t = 0, 1, 2, ...$$

L'égalité de cette expression avec l'expression en terme de durée effectivement observée ci-dessus est assurée sous l'hypothèse d'indépendance conditionnelle : $\{T_u^*(t_p), t_p = 0, 1, ...\} \perp T_p | X, V$ (indépendance entre les durées potentielles et la période effective de traitement, conditionnellement aux variables observées X et non observées V). Voir la preuve à l'Annexe A.

[11] Comme ci-dessus, on note que, $\forall t \geq 0$, l'événement ($T_u = t + t_p$) implique l'événement ($T_u \geq t_p$), et donc que l'intersection de ces deux événements est tout simplement l'événement ($T_u = t + t_p$).

De (4) et (5), on a encore :

$$\mathbb{P}[T_u = t + t_p | T_p = \infty, X, V] = h_u(t + t_p|\infty, X, V) S_u(t + t_p|\infty, X, V)$$

et

$$\mathbb{P}[T_u \geq t_p | T_p = \infty, X, V] = S_u(t_p|\infty, X, V)$$

de sorte qu'au total on a :

$$\begin{aligned} \bar{f}_u^R(t|t_p, X, V) &= \mathbb{P}\left[T_u^R = t | T_u \geq t_p, T_p = \infty, X, V\right], \quad \forall\, t = 0, 1, 2, ... \\ &= \frac{h_u(t + t_p|\infty, X, V) S_u(t + t_p|\infty, X, V)}{S_u(t_p|\infty, X, V)} \end{aligned} \tag{26}$$

où, par définition :

$$h_u(t|\infty, X, V) = \exp\left[\lambda_u(t) + \phi_u(X) + V_u\right]$$

et

$$S_u(t|\infty, X, V) = \begin{cases} 1 & , \text{si } t = 0 \\ \prod_{t^*=0}^{t-1} (1 - h_u(t^*|\infty, X, V)) & , \forall\, t = 1, 2, ... \end{cases}$$

La fonction de densité conditionnelle $\bar{f}_u^R(t|t_p, X, V)$ donne la probabilité que la durée 'résiduelle' T_u^R de chômage d'un individu pris au hasard soit (exactement) égale t, sachant qu'il n'a pas suivi de formation (i.e., $T_p = \infty$), qu'il était toujours au chômage en t_p (i.e., $T_u \geq t_p$), et ses caractéristiques observées X et non observées V. Elle équivaut à la distribution de probabilité la durée 'résiduelle' T_u^R de chômage qui aurait prévalu pour un individu pris au hasard parmi les individus qui ont suivi une formation en $T_p = t_p$, qui étaient toujours au chômage au début de cette formation ($T_u \geq t_p$), et qui ont des caractéristiques observées X et non observées V, s'il n'avait pas suivi de formation.

Les fonctions de densité conditionnelles $f_u^R(t|t_p, X, V)$ et $\bar{f}_u^R(t|t_p, X, V)$ permettent de contraster ce qui se passe lorsqu'un individu suit un formation en $T_p = t_p$ et ce qui se serait passé si le même individu n'avait pas suivi de formation, conditionnellement à ses caractéristiques observées X et à ses caractéristiques non observées V. Ces mêmes fonctions de densité, mais non conditionnelles à l'hétérogénéité non observée V, peuvent être obtenues en utilisant la distribution d'hétérogénéité conditionnelle $g^R(v|t_p, X)$ calculée précédemment[12]. Elles sont données par[13] :

[12] Pour rappel, $g^R(v|t_p, X)$ donne la distribution de l'hétérogénéité non observée parmi les individus qui ont suivi une formation en $T_p = t_p$, qui étaient toujours au chômage à cette date (i.e., $T_u \geq t_p$), et qui ont des caractéristiques observées X : $g^R(v|t_p, X) = \mathbb{P}[V = v | T_u \geq t_p, T_p = t_p, X]$.

[13] Exprimé en termes de durées potentielles $\{T_u^*(t_p), t_p = 0, 1, ...\}$, on a :

$$\begin{aligned} f_u^R(t|t_p, X) &= \mathbb{P}[T_u^*(t_p) = t + t_p | T_u^*(t_p) \geq t_p, T_p = t_p, X], \quad \forall\, t = 0, 1, 2, ... \\ &= \sum_{\{v\}} \mathbb{P}[T_u^*(t_p) = t + t_p | T_u^*(t_p) \geq t_p, T_p = t_p, X, V = v] \\ &\quad \times \mathbb{P}[V = v | T_u^*(t_p) \geq t_p, T_p = t_p, X] \end{aligned}$$

$$\begin{aligned} f_u^R(t|t_p,X) &= \sum_{\{v\}} f_u^R(t|t_p,X,v)g^R(v|t_p,X), \quad \forall\, t=0,1,2,... \\ &= \sum_{\{v\}} \frac{h_u(t+t_p|t_p,X,v)S_u(t+t_p|t_p,X,v)}{S_u(t_p|t_p,X,v)} g^R(v|t_p,X) \end{aligned}$$

et

$$\begin{aligned} \bar{f}_u^R(t|t_p,X) &= \sum_{\{v\}} \bar{f}_u^R(t|t_p,X,v)g^R(v|t_p,X), \quad \forall\, t=0,1,2,... \\ &= \sum_{\{v\}} \frac{h_u(t+t_p|\infty,X,v)S_u(t+t_p|\infty,X,v)}{S_u(t_p|\infty,X,v)} g^R(v|t_p,X) \end{aligned}$$

où, pour rappel :

$$g^R(v|t_p,X) = \frac{S_u(t_p|t_p,X,v)h_p(t_p|X,v)S_p(t_p|X,v)g(v)}{\sum_{\{v^*\}} S_u(t_p|t_p,X,v^*)h_p(t_p|X,v^*)S_p(t_p|X,v^*)g(v^*)}$$

On note que la même distribution d'hétérogénéité conditionnelle $g^R(v|t_p,X)$ est utilisée pour calculer la densité conditionnelle $f_u^R(t|t_p,X)$, i.e., ce qui se passe lorsqu'une formation est suivie, et son contrefactuel la densité conditionnelle $\bar{f}_u^R(t|t_p,X)$, i.e., ce qui se serait passé si aucune formation n'avait été suivie.

Pour contraster ce qui se passe pour un individu pris au hasard ayant des caractéristiques X lorsqu'il suit une formation en $T_p = t_p$, et ce qui se serait passé pour le même individu s'il n'avait suivi aucune formation, on peut comparer la médiane (ou tout autre quantile) des distributions $f_u^R(t|t_p,X)$ et $\bar{f}_u^R(t|t_p,X)$, pour toute valeur souhaitée de t_p et X. Une autre façon (complémentaire) de procéder est de comparer les fonctions de survie conditionnelles correspondantes aux fonctions de densités conditionnelles $f_u^R(t|t_p,X)$ et $\bar{f}_u^R(t|t_p,X)$. Ces fonctions de survie

---

et

$$\begin{aligned} \bar{f}_u^R(t|t_p,X) &= I\!P[T_u^*(\infty)=t+t_p|T_u^*(\infty)\geq t_p, T_p=t_p, X], \quad \forall\, t=0,1,2,... \\ &= \sum_{\{v\}} I\!P[T_u^*(\infty)=t+t_p|T_u^*(\infty)\geq t_p, T_p=t_p, X, V=v] \\ &\quad \times I\!P[V=v|T_u^*(\infty)\geq t_p, T_p=t_p, X] \end{aligned}$$

On note que lorsque $T_p = t_p$, on a $T_u = T_u^*(t_p)$, de sorte qu'on a :

$$I\!P[V=v|T_u^*(t_p)\geq t_p, T_p=t_p, X] = I\!P[V=v|T_u\geq t_p, T_p=t_p, X] = g^R(v|t_p,X)$$

On note par ailleurs que, sous l'hypothèse d'indépendance conditionnelle : $\{T_u^*(t_p), t_p=0,1,...\} \perp T_p\,|X,V$ et étant donné l'hypothèse de non anticipation qui implique en particulier que $\forall\, t<t_p$, $I\!P[T_u^*(t_p)=t|X,V]$ $= I\!P[T_u^*(\infty)=t|X,V]$, on a aussi (voir la preuve à l'Annexe B):

$$I\!P[V=v|T_u^*(\infty)\geq t_p, T_p=t_p, X] = I\!P[V=v|T_u\geq t_p, T_p=t_p, X] = g^R(v|t_p,X)$$

de sorte que c'est bien la même densité conditionnelle $g^R(v|t_p,X)$ de l'hétérogénéité non observée qui doit être utilisée pour obtenir $f_u^R(t|t_p,X)$ et $\bar{f}_u^R(t|t_p,X)$.

conditionnelles sont données par[14] :

$$S_u^R(t|t_p,X) = \begin{cases} 1 & , \text{si } t=0 \\ 1 - \sum\limits_{t^*=0}^{t-1} f_u^R(t^*|t_p,X) & , \forall\, t = 1,2,... \end{cases}$$

et

$$\bar{S}_u^R(t|t_p,X) = \begin{cases} 1 & , \text{si } t=0 \\ 1 - \sum\limits_{t^*=0}^{t-1} \bar{f}_u^R(t^*|t_p,X) & , \forall\, t = 1,2,... \end{cases}$$

et leur différence par[15] :

$$TT_u^R(t|t_p,X) = S_u^R(t|t_p,X) - \bar{S}_u^R(t|t_p,X)$$

$TT_u^R(t|t_p,X)$ s'interpréte comme l'effet du traitement, en termes de fonctions de survie, pour un individu pris au hasard traité en $T_p = t_p$ et ayant des caratéristiques observées X : pour chaque valeur de t, $TT_u^R(t|t_p,X)$ correspond à un classique 'effet moyen du traitement pour un individu traité' (en $T_p = t_p$ et ayant des caratéristiques caratéristiques observées X)[16]. On notera qu'étant donné que $TT_u^R(t|t_p,X)$ implique des fonctions de survie, un valeur positive de $TT_u^R(t|t_p,X)$ en t signifie un allongement (et non un raccourcissement) de la durée 'résiduelle' T_u^R de chômage par rapport à la période t, et donc un effet net globalement négatif de la formation jusqu'à t. A l'opposé, une valeur négative de $TT_u^R(t|t_p,X)$ en t signifie un raccourcissement de la durée 'résiduelle' T_u^R de chômage par rapport à la période t, et donc un effet net globalement positif de la formation à partir de t (si cette valeur négative de $TT_u^R(t|t_p,X)$ se maintient lorsque $t \to \infty$).

Les indicateurs décrits ci-dessus (médiane – ou autre quantile – de $f_u^R(t|t_p,X)$ et $\bar{f}_u^R(t|t_p,X)$, et $TT_u^R(t|t_p,X)$) mesurent l'effet du traitement pour un individu pris au hasard parmi les individus ayant suivi une formation $T_p = t_p$ et ayant des caractéristiques X. On peut définir l'effet agrégé du traitement pour l'ensemble – ou un sous-ensemble – des individus traités de l'échantillon, comme l'effet du traitement pour un individu pris au hasard parmi l'ensemble – ou un sous-ensemble

[14] Exprimé en termes de durées potentielles $\{T_u^*(t_p), t_p = 0,1,...\}$, on a :

$$S_u^R(t|t_p,X) = I\!P[T_u^*(t_p) \geq t + t_p | T_u^*(t_p) \geq t_p, T_p = t_p, X], \quad \forall\, t = 0,1,2,...$$

et

$$\bar{S}_u^R(t|t_p,X) = I\!P[T_u^*(\infty) \geq t + t_p | T_u^*(\infty) \geq t_p, T_p = t_p, X], \quad \forall\, t = 0,1,2,...$$

[15] Exprimé en termes de durées potentielles $\{T_u^*(t_p), t_p = 0,1,...\}$, on a :

$$\begin{aligned} & TT_u^R(t|t_p,X) \\ = \; & I\!P[T_u^*(t_p) \geq t + t_p | T_u^*(t_p) \geq t_p, T_p = t_p, X] - I\!P[T_u^*(\infty) \geq t + t_p | T_u^*(\infty) \geq t_p, T_p = t_p, X] \\ = \; & E[I(T_u^*(t_p) \geq t + t_p) | T_u^*(t_p) \geq t_p, T_p = t_p, X] - E[I(T_u^*(\infty) \geq t + t_p) | T_u^*(\infty) \geq t_p, T_p = t_p, X] \end{aligned}$$

[16] Cf. la seconde égalité de l'expression de $TT_u^R(t|t_p,X)$ en termes de durées potentielles dans la note en bas de page 15 ci-dessus.

– des individus traités. Soit $\mathcal{P} = \{(t_{p_i}, x_i),\ i = 1, ..., m\}$ l'ensemble des valeurs observées de (T_p, X) pour l'ensemble – ou un sous-ensemble – des individus traités de l'échantillon. On définit[17] :

$$f_u^{R,Ag}(t|\mathcal{P}) = \frac{1}{m}\sum_{i=1}^{m} f_u^R(t|t_{p_i}, x_i), \quad \forall\, t = 0, 1, 2, ...$$

et

$$\bar{f}_u^{R,Ag}(t|\mathcal{P}) = \frac{1}{m}\sum_{i=1}^{m} \bar{f}_u^R(t|t_{p_i}, x_i), \quad \forall\, t = 0, 1, 2, ...$$

La fonction de densité conditionnelle agrégée $f_u^{R,Ag}(t|\mathcal{P})$ donne la probabilité que la durée 'résiduelle' $T_u^R = T_u - T_p$ de chômage soient (exactement) égale t pour un individu pris au hasard parmi l'ensemble – ou un sous-ensemble – $\mathcal{P}$ des individus traités de l'échantillon, individus étaient toujours au chômage au début de leur formation (i.e., $T_u \geq T_p$) et dont les valeurs observées de (T_p, X) sont $\{(t_{p_i}, x_i),\ i = 1, ..., m\}$. Son contrefactuel $\bar{f}_u^{R,Ag}(t|\mathcal{P})$ donne ce qu'aurait été, pour un même individu, cette probabilité s'il n'avait pas suivi de formation. En termes de population, $f_u^{R,Ag}(t|\mathcal{P})$ donne la distribution attendue de fréquence des durées 'résiduelles' de chômage de l'ensemble – ou un sous-ensemble – $\mathcal{P}$ des individus traités de l'échantillon, et son contrefactuel ce qu'aurait été, pour le même ensemble $\mathcal{P}$ d'individus, cette distribution attendue de fréquence si ceux-ci n'avaient pas suivi de formation.

Comme ci-dessus, pour contraster ce qui se passe de façon agrégée pour les individus traités, et ce qui se serait passé de façon agrégée si ces mêmes individus n'avaient pas suivi de formation, on peut comparer la médiane (ou tout autre quantile) des distributions $f_u^{R,Ag}(t|\mathcal{P})$ et $\bar{f}_u^{R,Ag}(t|\mathcal{P})$, et encore calculer la différence des fonctions de survie qui y correspondent[18] :

$$TT_u^{R,Ag}(t|\mathcal{P}) = S_u^{R,Ag}(t|\mathcal{P}) - \bar{S}_u^{R,Ag}(t|\mathcal{P})$$

---

[17] Exprimé en termes de durées potentielles $\{T_u^*(t_p), t_p = 0, 1, ...\}$, on a :

$$\begin{aligned} f_u^{R,Ag}(t|\mathcal{P}) &= I\!P[T_u^*(T_p) = t + T_p | T_u^*(T_p) \geq T_p, \mathcal{P}], \quad \forall\, t = 0, 1, 2, ... \\ &= \frac{1}{m}\sum_{i=1}^{m} I\!P[T_u^*(t_{p_i}) = t + t_{p_i} | T_u^*(t_{p_i}) \geq t_{p_i}, T_p = t_{p_i}, X = x_i] \end{aligned}$$

et

$$\begin{aligned} \bar{f}_u^{R,Ag}(t|\mathcal{P}) &= I\!P[T_u^*(\infty) = t + T_p | T_u^*(\infty) \geq T_p, \mathcal{P}], \quad \forall\, t = 0, 1, 2, ... \\ &= \frac{1}{m}\sum_{i=1}^{m} I\!P[T_u^*(\infty) = t + t_{p_i} | T_u^*(\infty) \geq t_{p_i}, T_p = t_{p_i}, X = x_i] \end{aligned}$$

[18] Exprimé en termes de durées potentielles $\{T_u^*(t_p), t_p = 0, 1, ...\}$, on a :

$$\begin{aligned} & TT_u^{R,Ag}(t|\mathcal{P}) \\ = \;& I\!P[T_u^*(T_p) \geq t + T_p | T_u^*(T_p) \geq T_p, \mathcal{P}] - I\!P[T_u^*(\infty) \geq t + T_p | T_u^*(\infty) \geq T_p, \mathcal{P}] \\ = \;& E[I(T_u^*(T_p) \geq t + T_p) | T_u^*(T_p) \geq T_p, \mathcal{P}] - E[I(T_u^*(\infty) \geq t + T_p) | T_u^*(\infty) \geq T_p, \mathcal{P}] \end{aligned}$$

où :

$$S_u^{R,Ag}(t|\mathcal{P}) = \begin{cases} 1 & , \text{si } t = 0 \\ 1 - \sum\limits_{t^*=0}^{t-1} f_u^{R,Ag}(t^*|\mathcal{P}) & , \forall\, t = 1, 2, ... \end{cases}$$

et

$$\bar{S}_u^{R,Ag}(t|\mathcal{P}) = \begin{cases} 1 & , \text{si } t = 0 \\ 1 - \sum\limits_{t^*=0}^{t-1} \bar{f}_u^{R,Ag}(t^*|\mathcal{P}) & , \forall\, t = 1, 2, ... \end{cases}$$

$TT_u^{R,Ag}(t|\mathcal{P})$ s'interpréte comme l'effet agrégé du traitement, en termes de fonctions de survie, pour l'ensemble – ou un sous-ensemble – $\mathcal{P}$ des individus traités de l'échantillon : pour chaque valeur de t, $TT_u^{R,Ag}(t|\mathcal{P})$ correspond a un classique 'effet moyen du traitement pour les individus traités'. Son signe (positif ou négatif en fonction de t) s'interprète comme celui de $TT_u^R(t|t_p, X)$.

Les différentes quantités présentées ci-dessus sont des quantités théoriques. Elles peuvent simplement être estimées en évaluant les différents éléments qui les composent – i.e., les fonctions de hasard conditionnelles $h_u(t|t_p, X, V; \beta_u)$ et $h_p(t|X, V; \beta_p)$, et la fonction de densité $g(v; \alpha)$ de l'hétérogénéité non observée – aux paramètres estimés $\left(\hat{\beta}_u, \hat{\beta}_p, \hat{\alpha}\right)$ du modèle. Des écart-types et intervalles de confiance pour ces quantités peuvent par ailleurs être obtenus par la 'méthode delta'.

Annexe A

Dans cette annexe, on montre que sous l'hypothèse d'indépendance conditionnelle $\{T_u^*(t_p), t_p = 0, 1, ...\} \perp T_p\, |X, V$, on a bien :

$$\begin{aligned} \bar{f}_u^R(t|t_p, X, V) &= I\!P[T_u^*(\infty) = t + t_p | T_u^*(\infty) \geq t_p, T_p = t_p, X, V], \quad \forall\, t = 0, 1, 2, ... \\ &= I\!P[T_u = t + t_p | T_u \geq t_p, T_p = \infty, X, V] \end{aligned}$$

On a :

$$\begin{aligned} \bar{f}_u^R(t|t_p, X, V) &= I\!P[T_u^*(\infty) = t + t_p | T_u^*(\infty) \geq t_p, T_p = t_p, X, V] \\ &= \frac{I\!P[T_u^*(\infty) = t + t_p, T_u^*(\infty) \geq t_p | T_p = t_p, X, V]}{I\!P[T_u^*(\infty) \geq t_p | T_p = t_p, X, V]} \\ &= \frac{I\!P[T_u^*(\infty) = t + t_p | T_p = t_p, X, V]}{I\!P[T_u^*(\infty) \geq t_p | T_p = t_p, X, V]} \end{aligned}$$

où la seconde égalité vient du fait que l'événement $(T_u^*(\infty) = t + t_p)$ implique l'événement $(T_u^*(\infty) \geq t_p)$. Sous l'hypothèse d'indépendance conditionnelle, on a :

$$\begin{aligned} I\!P[T_u^*(\infty) = t + t_p | T_p = t_p, X, V] &= I\!P[T_u^*(\infty) = t + t_p | T_p = \infty, X, V] \\ &= I\!P[T_u = t + t_p | T_p = \infty, X, V] \end{aligned}$$

et

$$
\begin{aligned}
I\!\!P[T_u^*(\infty) \geq t_p | T_p = t_p, X, V] &= I\!\!P[T_u^*(\infty) \geq t_p | T_p = \infty, X, V] \\
&= I\!\!P[T_u \geq t_p | T_p = \infty, X, V]
\end{aligned}
$$

où la seconde égalité découle à chaque fois du fait que lorsque $T_p = \infty$, $T_u = T_u^*(\infty)$, de sorte qu'au total on a bien :

$$
\begin{aligned}
\bar{f}_u^R(t|t_p, X, V) &= I\!\!P[T_u^*(\infty) = t + t_p | T_u^*(\infty) \geq t_p, T_p = t_p, X, V] \\
&= \frac{I\!\!P[T_u^*(\infty) = t + t_p | T_p = t_p, X, V]}{I\!\!P[T_u^*(\infty) \geq t_p | T_p = t_p, X, V]} \\
&= \frac{I\!\!P[T_u = t + t_p | T_p = \infty, X, V]}{I\!\!P[T_u \geq t_p | T_p = \infty, X, V]} \\
&= \frac{I\!\!P[T_u = t + t_p, T_u \geq t_p | T_p = \infty, X, V]}{I\!\!P[T_u \geq t_p | T_p = \infty, X, V]} \\
&= I\!\!P[T_u = t + t_p | T_u \geq t_p, T_p = \infty, X, V]
\end{aligned}
$$

où l'avant-dernière égalité vient du fait que l'événement $(T_u = t + t_p)$ implique l'événement $(T_u \geq t_p)$.

Annexe B

Dans cette annexe, on montre que sous l'hypothèse d'indépendance conditionnelle : $\{T_u^*(t_p), t_p = 0, 1, ...\} \perp T_p | X, V$ et étant donné l'hypothèse de non anticipation qui implique en particulier que $\forall t < t_p, I\!\!P[T_u^*(t_p) = t | X, V] = I\!\!P[T_u^*(\infty) = t | X, V]$, on a bien :

$$
I\!\!P[V = v | T_u^*(\infty) \geq t_p, T_p = t_p, X] = I\!\!P[V = v | T_u \geq t_p, T_p = t_p, X] = g^R(v|t_p, X)
$$

Du théorème de Bayes, on a :

$$
\begin{aligned}
& I\!\!P[V = v | T_u^*(\infty) \geq t_p, T_p = t_p, X] \\
= & \frac{I\!\!P[T_u^*(\infty) \geq t_p, T_p = t_p | X, V = v] \, I\!\!P[V = v | X]}{I\!\!P[T_u^*(\infty) \geq t_p, T_p = t_p | X]} \\
= & \frac{I\!\!P[T_u^*(\infty) \geq t_p, T_p = t_p | X, V = v] \, I\!\!P[V = v | X]}{\sum\limits_{\{v^*\}} I\!\!P[T_u^*(\infty) \geq t_p, T_p = t_p | X, V = v^*] \, I\!\!P[V = v^* | X]}
\end{aligned}
$$

Sous l'hypothèse d'indépendance conditionnelle, on a :

$$
\begin{aligned}
& I\!\!P[T_u^*(\infty) \geq t_p, T_p = t_p | X, V = v] \\
= & I\!\!P[T_u^*(\infty) \geq t_p | X, V = v] \, I\!\!P[T_p = t_p | X, V = v] \\
= & I\!\!P[T_u^*(t_p) \geq t_p | X, V = v] \, I\!\!P[T_p = t_p | X, V = v] \\
= & I\!\!P[T_u^*(t_p) \geq t_p | T_p = t_p, X, V = v] \, I\!\!P[T_p = t_p | X, V = v] \\
= & I\!\!P[T_u \geq t_p | T_p = t_p, X, V = v] \, I\!\!P[T_p = t_p | X, V = v] \\
= & I\!\!P[T_u \geq t_p, T_p = t_p | X, V = v]
\end{aligned}
$$

où la seconde égalité découle de l'hypothèse de non anticipation et l'avant-dernière égalité du fait que lorsque $T_p = t_p$, $T_u = T_u^*(t_p)$, de sorte qu'au total on a bien :

$$\begin{aligned}
& I\!P[V = v | T_u^*(\infty) \geq t_p, T_p = t_p, X] \\
= \quad & \frac{I\!P[T_u^*(\infty) \geq t_p, T_p = t_p | X, V = v]\, I\!P[V = v|X]}{\sum_{\{v^*\}} I\!P[T_u^*(\infty) \geq t_p, T_p = t_p | X, V = v^*]\, I\!P[V = v^*|X]} \\
= \quad & \frac{I\!P[T_u \geq t_p, T_p = t_p | X, V = v]\, I\!P[V = v|X]}{\sum_{\{v^*\}} I\!P[T_u \geq t_p, T_p = t_p | X, V = v^*]\, I\!P[V = v^*|X]} \\
= \quad & \frac{I\!P[T_u \geq t_p, T_p = t_p | X, V = v]\, I\!P[V = v|X]}{I\!P[T_u \geq t_p, T_p = t_p | X]} \\
= \quad & I\!P[V = v | T_u \geq t_p, T_p = t_p, X] = g^R(v | t_p, X)
\end{aligned}$$

où la dernière égalité découle à nouveau du théorème de Bayes.