Transformations robustes sans biais en présence de censure

Damien Bousquet

Institut de Recherche Mathématique de Rennes
Institut National des Sciences Appliquées de Rennes

Damien.Bousquet@insa-rennes.fr

Résumé

Notre problème est de construire des transformations sans biais et de variance minimale d'une fonction quelconque de la variable aléatoire non observée dans les échantillons censurés. Nous supposons un mécanisme indépendant de censure aléatoire à droite. Quand la fonction cible a une seule variable réelle, ce problème a été résolu par l'auteur dans [7]. Notre contribution [1] est de montrer que ce problème a une solution dans le cas général, quand la fonction cible a plusieurs variables. Nous donnons sa forme explicite. La transformation solution de notre problème a une autre propriété remarquable. Son expression fait intervenir la fonction de survie de la variable non observée et celle de la variable de censure. Cette transformation est doublement robuste. En effet, il suffit seulement que la survie de la variable non observée ou celle de la censure soient données, pour avoir la propriété de non biais. Pour une fonction cible égale à l'identité des réels, la solution a été présentée par les auteurs dans [6].

Mots clés

biais, censure, robustesse, variance

Abstract

Our study aims at building uniformly minimum variance unbiased estimators of a real valued function from several variables that are the non observed data of a censored sample. We suppose a right independent censoring mechanism. When this function has only one variable, this problem has been solved by the author in [7]. Our contribution [1] is to show that there exists a solution in the general case, when the baseline function has more than one variable. The solution is explicit. This solution has another interesting property. The formula depends on the survival from the non observed random variable and the survival from the censoring random variable. This transformation is doubly robust.

Indeed, it is sufficient to give only the survival from the non observed random variable or the survival from the censoring random variable, to obtain an unbiased transformation from the quantity of interest. When the map is the identity from the real numbers, this solution has been presented by the authors in [6].

Keywords

bias, censoring, robustness, variance

1 Introduction

Dans ce travail, nous construisons des transformations sans biais d'une fonction cible du vecteur aléatoire d'intérêt dans un modèle de censure aléatoire indépendante à droite [4]. Notre travail [1] est une généralisation de la transformation Inverse Probability of Censoring Weighted [3]. Notre approche est exhaustive. Nous déterminons toutes les transformations sans biais possibles sous conditions d'intégrabilité des quantités en jeu. La fonction cible est à valeurs dans les réels et elle a n variables réelles, quand l'échantillon de base comporte n observations. Ici, l'entier n est arbitraire. La fonction cible étant fixée, nous déterminons toutes les transformations sans biais possibles. Notre critère d'optimalité est un critère de variance. Le calcul montre qu'il existe une et une seule transformation de variance minimale. Cette transformation optimale a une propriété remarquable. Elle est doublement robuste. En effet, elle dépend de deux quantités théoriques, la survie de la variable aléatoire d'intérêt, et de la survie de la variable aléatoire de censure. Il suffit qu'une seule parmi les deux, soit exactement spécifiée, pour que la propriété de non biais soit vérifiée. Dans le cas des échantillons d'une observation $n = 1$, cette transformation est présentée par l'auteur dans [7], qui met en évidence sa variance minimale et, par les auteurs dans [6] qui notent son caractère doublement robuste. Notre contribution [1] est de généraliser aux échantillons de n observations, n arbitraire. Par exemple, pour estimer sans biais la variance, cela nécessite de manière générale au moins deux observations [5].

2 Transformations sans biais

Nous présentons pour simplifier notre théorie dans le cas d'une observation $n = 1$. Ce raisonnement se généralise alors aux échantillons de n observations, n arbitraire. Le cas général est traité par l'auteur dans [1]. Nous unifions à présent les résultats obtenus par les auteurs dans [6] et [7]. Notre approche est nouvelle.

Dans la suite, chaque expression ou résultat sont donnés sous réserve d'intégrabilité des quantités en jeu, ce que nous omettrons de répéter.

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, $Y : \Omega \to \mathbf{R}$ et $C : \Omega \to \mathbf{R}$ des variables aléatoires indépendantes, $T = \min(Y, C)$, $D = 1_{Y \leq C}$, $\gamma : \mathbf{R} \to \mathbf{R}$ une fonction borélienne fixée, $V_1 = \{(t_1, t_2) \in \mathbf{R}^2 \ ; \ t_1 \leq t_2\}$ et $V_0 = \mathbf{R}^2 - V_1$.

Nous notons $S_Y : \mathbf{R} \to [0, 1]$ la fonction de survie de la variable aléatoire non observée $Y : \Omega \to \mathbf{R}$, et $S_C : \mathbf{R} \to [0, 1]$ celle de la variable aléatoire de censure $C : \Omega \to \mathbf{R}$.

Nous supposons que pour tout réel c,

$$S_C(c) = \mathbb{P}(C > c) > 0$$

Nous donnons maintenant la transformation générique sans biais $\mathrm{U}_{C,\gamma,\theta_0} : \Omega \to \mathbf{R}$.

Soient $\theta_0 : \mathbf{R} \to \mathbf{R}$ une fonction borélienne et pour tout réel y,

$$\theta_1(y) = \frac{1}{S_C(y)}[\gamma(y) - \int_{\mathbf{R}} 1_{V_0}(y, c)\theta_0(c)d\mathbb{P}_C(c)]$$

et pour tout $(t, d) \in \mathbf{R} \times \{0, 1\}$,

$$\mathrm{U}_{C,\gamma,\theta_0}(t, d) = (1 - d)\theta_0(t) + d\theta_1(t)$$

Proposition 2.1 *Nous avons alors,*

$$\mathbb{E}[\mathrm{U}_{C,\gamma,\theta_0}(T, D)] = \mathbb{E}[\gamma \circ Y]$$

Nous avons décrit toutes les transformations sans biais possibles, les paramètres libres étant les fonctions γ et θ_0.

Pour $\theta_0 = 0$, nous retrouvons la transformation Inverse Probability of Censoring Weighted standard.

Nous notons $\mathrm{C}(\mathbf{R}, \mathbf{R})$ l'ensemble de toutes les fonctions continues de $\mathbf{R}$ dans $\mathbf{R}$ et Θ l'ensemble de toutes les fonctions continues $\theta_0 : \mathbf{R} \to \mathbf{R}$ telles que pour tout réel y,

$$\int_{\mathbf{R}^2} 1_{V_0}(y, c)\theta_0^2(c)d\mathbb{P}_{(Y,C)}(y, c) \in \mathbf{R}_{\geq 0}$$

$$\int_{\mathbf{R}} \frac{1}{S_C(y)}\{\int_{\mathbf{R}} 1_{V_0}(y, c)\theta_0(c)d\mathbb{P}_C(c)\}^2 d\mathbb{P}_Y(y) \in \mathbf{R}_{\geq 0}$$

$$\int_{\mathbf{R}^2} 1_{V_1}(y, c)\theta_1^2(y)d\mathbb{P}_{(Y,C)}(y, c) \in \mathbf{R}_{\geq 0}$$

Si $\theta_0 \in \Theta$, nous avons,

$$\mathbb{E}[\mathrm{U}^2_{C,\gamma,\theta_0}(T, D)] \in \mathbf{R}_{\geq 0}$$

Proposition 2.2 *Nous supposons que la fonction* $y \mapsto \frac{\gamma^2(y)}{S_C(y)}$ *de* $\mathbf{R}$ *dans* $\mathbf{R}_{\geq 0}$ *est intégrable pour la mesure* $\mathbb{P}_Y$*, de sorte que la fonction nulle est dans* Θ*. Alors le sous ensemble* Θ *de* $\mathrm{C}(\mathbf{R}, \mathbf{R})$ *est un sous espace vectoriel de* $\mathrm{C}(\mathbf{R}, \mathbf{R})$*.*

Nous introduisons notre application d'intérêt χ, pour tout $\theta_0 \in \Theta$,

$$\chi(\theta_0) = \mathbb{E}[\mathrm{U}^2_{C,\gamma,\theta_0}(T, D)]$$

et une norme N sur l'espace vectoriel Θ,

$$\mathrm{N}(\theta_0) = [\int_{\mathbf{R}} \frac{1}{S_C(y)} \{\int_{\mathbf{R}} 1_{V_0}(y,c)\theta_0(c) d\mathbb{P}_C(c)\}^2 d\mathbb{P}_Y(y) + \int_{\mathbf{R}^2} 1_{V_0}(y,c)\theta_0^2(c) d\mathbb{P}_{(Y,C)}(y,c)]^{\frac{1}{2}}$$

Nous supposons l'existence de deux fonctions $f_Y : \mathbf{R} \to \mathbf{R}_{>0}$ et $f_C : \mathbf{R} \to \mathbf{R}_{>0}$ qui sont intégrables sur $\mathbf{R}$ pour la mesure de Lebesgue et telles que pour tout $B \in B(\mathbf{R})$, nous avons,

$$\mathbb{P}_Y(B) = \int_{\mathbf{R}} 1_B(t) f_Y(t) d\lambda(t)$$

$$\mathbb{P}_C(B) = \int_{\mathbf{R}} 1_B(t) f_C(t) d\lambda(t)$$

Sous cette hypothèse, nous avons pour tout réel t,

$$S_Y(t) = \mathbb{P}(Y > t) > 0$$

$$S_C(t) = \mathbb{P}(C > t) > 0$$

Proposition 2.3 *L'application* N *de* Θ *dans* $\mathbf{R}_{\geq 0}$ *définit alors une norme sur l'espace vectoriel* Θ*. De plus, l'application* χ *de* Θ *dans* $\mathbf{R}_{\geq 0}$ *est fortement convexe de module* 2 *sur* Θ *pour la norme* N*.*

Notre but est de minimiser cette fonction χ sur Θ. Cette fonction est strictement convexe.

Proposition 2.4 *L'application* χ *est différentiable sur* Θ *pour la norme* N *et admet un point critique* $\theta_0^0 \in \Theta$*.*

L'application χ est minimale à ce point. Le point θ_0^0 fournit notre transformation optimale.

Nous donnons maintenant sa forme explicite.

Soit $\eta : \mathbf{R} \to \mathbf{R}$ la fonction telle que pour tout réel c,

$$\eta(c) = \frac{1}{S_Y(c)} \int_{\mathbf{R}} 1_{V_0}(y, c)\gamma(y) d\mathbb{P}_Y(y)$$

et pour tout réel y, la fonction μ suivante,

$$\mu(y) = \frac{\gamma(y)}{S_C(y)} - \int_{\mathbf{R}} 1_{V_1}(c, y) \frac{\eta(c)}{S_C^2(c)} d\mathbb{P}_C(c)$$

Proposition 2.5 *Pour tout réel c, nous avons,*

$$\theta_0^0(c) = \frac{1}{S_Y(c)} \int_{\mathbf{R}} 1_{V_0}(y, c)\mu(y) d\mathbb{P}_Y(y)$$

Nous faisons l'hypothèse suivante,

$$\int_{\mathbf{R}} \frac{d\mathbb{P}_Y(y)}{S_C(y)} \in \mathbf{R}_{\geq 0}$$

$$\lim_{t \to +\infty} \frac{S_Y(t)}{S_C(t)} = 0$$

Pour tout réel y, nous avons alors,

$$\theta_1^0(y) = \mu(y)$$

Le théorème suivant montre que cette transformation est de variance minimale.

Théorème 2.1 *Si $\theta_0^0 \in \Theta$, nous avons pour tout $\theta_0 \in \Theta$ tel que $\theta_0 \neq \theta_0^0$,*

$$\mathbb{V}[\mathrm{U}_{C,\gamma,\theta_0}(T, D)] > \mathbb{V}[\mathrm{U}_{C,\gamma,\theta_0^0}(T, D)]$$

Nous montrons que la transformation optimale est aussi robuste.

Soient $Y_0 : \Omega \to \mathbf{R}$ et $C_0 : \Omega \to \mathbf{R}$ des variables aléatoires indépendantes et,

$$T_0 = \min(Y_0, C_0)$$

$$D_0 = 1_{Y_0 \leq C_0}$$

Théorème 2.2 *Si nous avons $S_C = S_{C_0}$ ou bien $S_Y = S_{Y_0}$, alors nous avons toujours,*

$$\mathbb{E}[\mathrm{U}_{C,\gamma,\theta_0^0}(T_0, D_0)] = \mathbb{E}[\gamma(Y_0)]$$

Quand $\gamma = \mathrm{id}_{\mathbf{R}}$, nous retrouvons la transformation proposée par les auteurs dans [6].

3 Conclusion

Nous avons ainsi construit une transformation non biaisée d'une fonction quelconque de la variable aléatoire d'intérêt en présence de censure. Cette transformation est de variance minimale et robuste. Toutefois, ces transformations ne sont pas des estimateurs au sens statistique, parce qu'ils dépendent de quantités théoriques non observées. Nous pouvons substituer par exemple aux fonctions de survie théoriques leurs estimateurs de Kaplan et Meier [2]. L'estimateur obtenu ne sera plus a priori sans biais. Cependant, la propriété de robustesse de notre transformation théorique et sa variance minimale, nous laissent espérer qu'elle peut être le point de départ pour construire de bons estimateurs de quantités générales avec censure.

Références

[1] Bousquet, D. (2013) *Robust unbiased transformations with censoring* hal-00825893

[2] Kaplan, E. L. et Meier, P. (1958) *Nonparametric estimation from incomplete observations* Journal of The American Statistical Association, 53, 282, 457–481

[3] Koul, H., Susarla, V. et Van Ryzin, J. (1981) *Regression analysis with randomly right-censored data* The Annals of Statistics, 9, 6, 1276–1288

[4] Lawless, J. F. (2003) *Statistical Models and Methods for Lifetime Data* Wiley Series in Probability and Statistics

[5] Lehmann, E. L. (1983) *Theory of Point Estimation* A Wiley Publication in Mathematical Statistics

[6] Rubin, D. B. et Van Der Laan, M. J. (2007) *A doubly robust censoring unbiased transformation* The International Journal of Biostatistics, 3, 1, Article quatre

[7] Suzukawa, A. (2004) *Unbiased estimation of functionals under random censorship* Journal of The Japan Statistical Society, 34, 2, 153–172

[8] Zheng, Z. (1987) *A class of estimators of the parameters in linear regression with censored data* Acta Mathematicae Applicatae Sinica, 3, 3, 231–241