

Utilisation du *bootstrap* pour les problèmes statistiques liés à l'estimation des paramètres

Rudy Palm

Unité de Statistique et Informatique. Faculté universitaire des Sciences agronomiques de Gembloux. Avenue de la Faculté d'Agronomie 8. B-5030 Gembloux (Belgique). E-mail : palm.r@fsagx.ac.be

Reçu le 7 mars 2002, accepté le 29 avril 2002.

Cet article décrit l'utilisation du *bootstrap* pour le calcul de l'erreur-standard et du biais d'un estimateur et pour la détermination des limites de confiance d'un paramètre estimé. Les différentes méthodes exposées sont illustrées par un exemple.

Mots-clés. *Bootstrap*, erreur-standard, biais, intervalle de confiance, jackknife, moyenne, médiane, variance, statistique.

Use of bootstrap for statistical problems related to estimation of parameters. This paper describes bootstrapping for the estimation of the standard error and the bias of an estimator and for computing confidence intervals. The methods are illustrated by an example.

Keywords. Bootstrap, standard error, bias, confidence interval, jackknife, mean, median, variance, statistics.

1. INTRODUCTION

Les méthodes classiques d'inférence statistique ne permettent pas d'obtenir des réponses correctes à tous les problèmes concrets que se pose l'utilisateur. Elles ne sont en effet valables que sous des conditions d'application particulières. Ainsi, par exemple, le test *t* de Student d'égalité des moyennes suppose que les deux populations-parents sont normales, de même variance et que les deux échantillons sont aléatoires, simples et indépendants. Le calcul de l'intervalle de confiance d'une variance par l'intermédiaire des variables χ^2 suppose que la population-parent est normale et que l'échantillon est aléatoire et simple. L'inférence statistique classique en régression suppose, outre les conditions d'application relatives à la population et à l'échantillon, que le modèle est ajusté au sens des moindres carrés.

Que peut faire l'utilisateur, en pratique, lorsque ces conditions d'application ne sont pas remplies ? Différentes attitudes sont possibles.

Dans certains cas, les méthodes classiques sont utilisées malgré le non-respect des conditions. Cette utilisation est alors justifiée par le caractère robuste des méthodes qui garantissent que les résultats de l'inférence restent approximativement valables. C'est par exemple le cas pour la comparaison de moyennes de populations qui s'écartent modérément des distributions normales, lorsque les effectifs sont suffisamment grands.

Le recours à des transformations de variables permet, dans certains cas, de se rapprocher des conditions d'application. Ainsi une transformation logarithmique, par exemple, peut rendre normales des distributions qui, au départ, ne le sont pas.

Une troisième attitude consiste à abandonner les méthodes paramétriques d'inférence statistique au profit de méthodes non paramétriques, pour lesquelles les conditions d'application sont bien moins restrictives. Pour le problème de la comparaison de deux moyennes évoqué ci-dessus, le test *t* de Student sera par exemple remplacé par le test des rangs de Wilcoxon, aussi appelé test de Mann-Whitney (Dagnelie, 1998).

Les solutions proposées ci-dessus permettent incontestablement d'élargir l'éventail des problèmes auxquels une solution peut être apportée. Elles ne permettent cependant pas de résoudre tous les problèmes. L'accès généralisé à des moyens de calcul puissants a permis le développement de méthodes d'inférence statistique basées sur l'utilisation intensive de l'ordinateur. Le *bootstrap* fait partie de ces méthodes.

Le mot *bootstrap* provient de l'expression anglaise "to pull oneself up by one's bootstrap" (Efron, Tibshirani, 1993), qui signifie littéralement "se soulever en tirant sur les languettes de ses bottes". Le mot *bootstrap* fait penser à des traductions telles que "à la force du poignet" ou "par soi-même" ou "passe partout" (Dagnelie, 1998), mais en fait il n'est jamais traduit dans la littérature scientifique d'expression française.

Le principe général de la méthode est de rééchantillonner un grand nombre de fois l'échantillon initial qui a été réellement prélevé dans la population, l'inférence statistique étant basée sur les résultats des échantillons ainsi obtenus.

L'objectif de cette note est de décrire comment le *bootstrap* peut être utilisé pour résoudre les problèmes d'inférence statistique en relation avec l'estimation des paramètres. Nous présentons d'abord les méthodes de rééchantillonnage. Ensuite, nous examinons l'estimation de l'erreur-standard et du biais d'un estimateur. Nous donnons alors quelques méthodes de détermination de l'intervalle de confiance d'un paramètre avant de conclure.

Des informations plus détaillées concernant le *bootstrap* et d'autres méthodes de rééchantillonnage sont données, notamment, par Chernick (1999), Efron et Tibshirani (1993) et Manly (1997).

2. MÉTHODES DE RÉÉCHANTILLONNAGE

2.1. Bootstrap des individus

On considère un échantillon de n observations :

$$x_1, x_2, \dots, x_i, \dots, x_n,$$

prélevé de manière aléatoire et simple dans une population. Ces observations peuvent concerner une seule variable, ou, au contraire, être relatives à plusieurs variables. Dans ce cas, les x_i représentent des vecteurs de dimension p , p étant le nombre de variables. Afin de ne pas alourdir les notations, nous ne distinguerons pas ces deux situations et, de manière plus condensée, nous désignerons l'échantillon initial par le symbole x , qu'il s'agisse d'un vecteur ou d'une matrice.

Le principe de la méthode du *bootstrap* est de prélever une série d'échantillons aléatoires et simples avec remise de n observations dans l'échantillon initial, considéré comme une population. Ces échantillons successifs seront notés :

$$x_1^*, x_2^*, \dots, x_k^*, \dots, x_B^*,$$

B étant le nombre de rééchantillonnages effectués.

À titre d'illustration, nous considérons le problème de l'estimation de diverses caractéristiques de la population des tailles des exploitations agricoles de la Région wallonne, à partir d'un échantillon aléatoire et simple de 100 observations.

La population a été simulée à partir de la distribution groupée résultant du recensement agricole et horticole au 15 mai 1995 (INS, 1996). Pour une classe donnée, d'effectif n_i et de limites de classe $x_{i \text{ inf}}$ et $x_{i \text{ sup}}$, on a généré n_i nombres aléatoires appartenant à une distribution uniforme dans le domaine $(x_{i \text{ inf}}, x_{i \text{ sup}})$. On a ainsi obtenu les tailles simulées des 24.719 exploitations. La **figure 1** reprend l'histogramme et

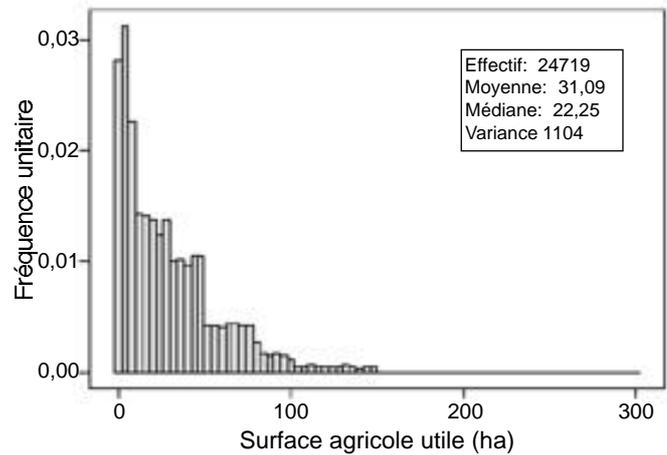


Figure 1. Population simulée des tailles des exploitations de la Région wallonne en 1995 — *Simulated population of the farms sizes in the Walloon Region in 1995.*

les principaux paramètres de cette population dont la caractéristique la plus marquante est la très forte dissymétrie gauche.

L'exemple présente donc un caractère artificiel, dans la mesure où on connaît exactement les caractéristiques de la population, celle-ci étant simulée. L'intérêt de disposer d'une population est de permettre une répétition des calculs, comme nous le verrons au paragraphe 4.7.

Dans la population théorique en question, on a sélectionné, de manière aléatoire et simple, un échantillon de 100 observations. La deuxième colonne du **tableau 1**, notée x , donne les premières et les dernières observations de l'échantillon, après classement des données par ordre croissant. Les trois colonnes suivantes donnent les premières et les dernières observations de trois échantillons de 100 observations prélevés dans l'échantillon initial et notés $x_1^*, x_2^*, \dots, x_3^*$, ceux-ci ayant également été classés par ordre croissant. On constate, par exemple, que pour l'échantillon x_1^* , la première observation de l'échantillon initial a été sélectionnée deux fois et que

Tableau 1. Échantillon initial x et résultats de trois rééchantillonnages, x_1^* , x_2^* et x_3^* (données partielles) — *Initial sample x and three bootstrap samples, x_1^* , x_2^* and x_3^* (partial data).*

Numéro d'ordre	x	x_1^*	x_2^*	x_3^*
1	0,00	0,00	0,00	0,00
2	0,18	0,00	0,00	0,00
3	0,36	0,36	0,18	0,36
4	0,46	1,04	0,18	0,36
...				
97	96,81	85,55	96,81	91,61
98	98,60	85,55	98,60	96,81
99	133,60	91,61	133,60	98,60
100	145,21	96,81	133,60	98,60

la deuxième observation de l'échantillon initial n'a, par contre, pas été sélectionnée. Le rééchantillonnage se faisant avec remise, il est tout à fait normal que certaines observations de l'échantillon initial soient absentes, ou au contraire apparaissent plus d'une fois.

Pour l'ensemble des B échantillons obtenus par *bootstrap* (*bootstrap sample*), les observations x_i n'apparaissent pas en nombre égal et on peut définir les proportions d'apparition P_i^* de chacune des observations, P_i^* étant égal au nombre de fois que l'observation x_i a été prélevée pour l'ensemble des B échantillons, divisé par le nombre total de prélèvements, qui est égal à nB . Ces proportions P_i^* interviennent dans certaines estimations (paragraphe 3.2). Des méthodes de rééchantillonnage assurant l'égalité de ces proportions sont également proposées. Cette approche porte le nom de rééchantillonnage balancé.

2.2. Bootstrap des résidus

La technique de rééchantillonnage présentée ci-dessus est la plus simple et la plus courante. Des méthodes un peu différentes sont utilisées pour des applications particulières. Ainsi, dans les problèmes de régression lorsque les valeurs des variables explicatives sont fixées *a priori* par l'utilisateur, le rééchantillonnage d'individus peut difficilement se justifier. Dans une telle situation, on peut remplacer le *bootstrap* des individus par le *bootstrap* des résidus.

Soit \mathbf{y} le vecteur de la variable à expliquer et \mathbf{Z} la matrice des variables explicatives. L'échantillon initial est donc décrit par la juxtaposition du vecteur \mathbf{y} et de la matrice \mathbf{Z} .

Soit $\hat{\mathbf{y}}$ le vecteur des coefficients estimés par une méthode donnée d'ajustement, qui n'est pas nécessairement la méthode des moindres carrés. On peut calculer le vecteur des valeurs estimées de la variable à expliquer et en déduire le vecteur des résidus :

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}.$$

Dans le cas du modèle linéaire, on a :

$$\hat{\mathbf{y}} = \mathbf{Z} \hat{\boldsymbol{\beta}}$$

mais, de manière plus générale, y_i peut être une fonction quelconque des valeurs observées des variables explicatives et des paramètres estimés :

$$\hat{y}_i = f(\mathbf{z}_i, \hat{\boldsymbol{\beta}}).$$

Soit \mathbf{e}_k^* , un échantillon aléatoire et simple prélevé avec remise dans le vecteur \mathbf{e} . Si le vecteur \mathbf{e} n'est pas de moyenne nulle, il est nécessaire de soustraire cette moyenne de chacun des résidus, avant de procéder au rééchantillonnage (Léger *et al.*, 1992).

En additionnant \mathbf{e}_k^* à la partie déterministe du modèle $f(\mathbf{Z}, \hat{\boldsymbol{\beta}})$, on obtient le vecteur \mathbf{y}_k^* :

$$\mathbf{y}_k^* = f(\mathbf{Z}, \hat{\boldsymbol{\beta}}) + \mathbf{e}_k^*.$$

La juxtaposition de \mathbf{y}_k^* et de \mathbf{Z} constitue le résultat du rééchantillonnage, \mathbf{x}_k^* . Comme précédemment, la procédure décrite est répétée B fois.

Il faut noter que l'application du *bootstrap* à la régression n'implique pas nécessairement le *bootstrap* des résidus. Le choix de l'une ou de l'autre méthode dépend du caractère fixe ou aléatoire de la matrice \mathbf{Z} , mais dépend aussi des hypothèses relatives au modèle. Le rééchantillonnage des résidus suppose en effet que les résidus ne sont pas fonction des variables explicatives, ce qui n'est, par exemple, pas le cas en présence d'inégalité des variances conditionnelles ou d'inadéquation de la relation.

3. ERREUR-STANDARD ET BIAIS D'UN PARAMÈTRE

3.1. Estimation de l'erreur-standard

Soit un paramètre θ de la population et soit :

$$\hat{\theta} = f(x_1, \dots, x_n) = f(\mathbf{x})$$

une estimation de ce paramètre, obtenue à partir des données de l'échantillon initial \mathbf{x} .

Chaque échantillon obtenu par rééchantillonnage permet de calculer une répétition du *bootstrap* (*bootstrap replication*) de l'estimation $\hat{\theta}$:

$$\hat{\theta}_k = f(\mathbf{x}_k^*) \quad (k = 1, \dots, B),$$

la fonction f étant la même que celle utilisée pour la définition de $\hat{\theta}$.

Considérons qu'on s'intéresse à la moyenne, à la médiane et à la variance de la distribution des tailles des exploitations agricoles et qu'on se propose d'estimer ces trois paramètres à partir de l'échantillon \mathbf{x} . Si on utilise les estimateurs classiques, le paramètre θ s'écrit, successivement pour les trois paramètres considérés :

$$\bar{x} = \frac{1}{100} \sum_{i=1}^{100} x_i, \quad \tilde{x} = \frac{1}{2} (x_{[50]} + x_{[51]})$$

$$\text{et } s^2 = \frac{1}{99} \sum_{i=1}^{100} (x_i - \bar{x})^2,$$

$x_{[50]}$ et $x_{[51]}$ étant les observations de rangs 50 et 51 de l'échantillon initial. Les valeurs numériques pour ces trois estimations sont données dans la première partie du **tableau 2**, sur la ligne intitulée $\hat{\theta}$.

Tableau 2. Paramètres estimés pour l'échantillon initial ($\hat{\lambda}$) et pour les trois premiers échantillons obtenus par rééchantillonnage ($\hat{\lambda}_1^*$, $\hat{\lambda}_2^*$ et $\hat{\lambda}_3^*$); moyennes ($\hat{\lambda}^*$) et écarts-types ($\hat{\lambda}^{\wedge}_*$) des paramètres estimés pour 1.000 rééchantillonnages – *Parameters estimated from the initial sample ($\hat{\lambda}$) and from the first three bootstrap samples ($\hat{\lambda}_1^*$, $\hat{\lambda}_2^*$ et $\hat{\lambda}_3^*$); means ($\hat{\lambda}^*$) and standard deviations ($\hat{\lambda}^{\wedge}_*$) for 1,000 bootstrap replicates.*

Paramètre	Moyenne	Médiane	Variance
$\hat{\lambda}$	28,13	21,56	854,63
$\hat{\lambda}_1^*$	27,84	18,91	667,19
$\hat{\lambda}_2^*$	26,32	19,95	796,93
$\hat{\lambda}_3^*$	31,22	23,15	708,62
.			
.			
.			
$\hat{\lambda}^*$	27,99	20,44	843,58
$\hat{\lambda}^{\wedge}_*$	2,89	2,53	184,25

Les calculs des trois paramètres peuvent être répétés pour les échantillons x_1^* , x_2^* et x_3^* . Les résultats obtenus sont repris dans la seconde partie du **tableau 2**, sur les lignes intitulées $\hat{\lambda}_1^*$, $\hat{\lambda}_2^*$ et $\hat{\lambda}_3^*$. Cette seconde partie du tableau peut évidemment être complétée au fur et à mesure des rééchantillonnages fournissant x_4^* , x_5^* , ..., x_B^* .

Disposant des B répétitions, on peut déterminer la moyenne :

$$\hat{\lambda} = \frac{1}{B} \sum_{k=1}^B \hat{\lambda}_k$$

et l'écart-type des $\hat{\lambda}_k^*$:

$$\hat{\lambda}^{\wedge}_* = \sqrt{\frac{B}{B-1} \sum_{k=1}^B (\hat{\lambda}_k^* - \hat{\lambda}^*)^2}$$

Pour l'exemple ci-dessus, 1.000 rééchantillonnages ont été réalisés. Les **figures 2, 3** et **4** donnent la distribution des valeurs obtenues pour les trois paramètres considérés. Pour la moyenne et pour la variance, on constate que la distribution est en cloche et relativement symétrique. Par contre, la distribution des médianes est assez différente puisqu'elle présente un caractère bimodal assez prononcé. La moyenne et l'écart-type des 1.000 moyennes, des 1.000 médianes et des 1.000 variances sont donnés dans la troisième partie du **tableau 2**, dans les lignes intitulées $\hat{\lambda}^*$ et $\hat{\lambda}^{\wedge}_*$.

L'écart-type $\hat{\lambda}^{\wedge}_*$ est une estimation de l'erreur-standard de l'estimateur du paramètre $\hat{\lambda}$. Pour les situations où on dispose d'un estimateur de cette erreur-standard, et pour autant que les conditions d'application soient remplies, on peut montrer que

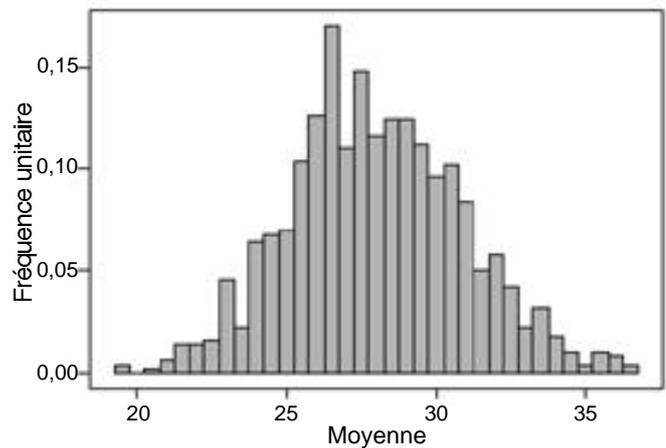


Figure 2. Distribution des moyennes de 1.000 échantillons obtenus par *bootstrap* — *Distribution of the means of 1,000 bootstrap samples.*

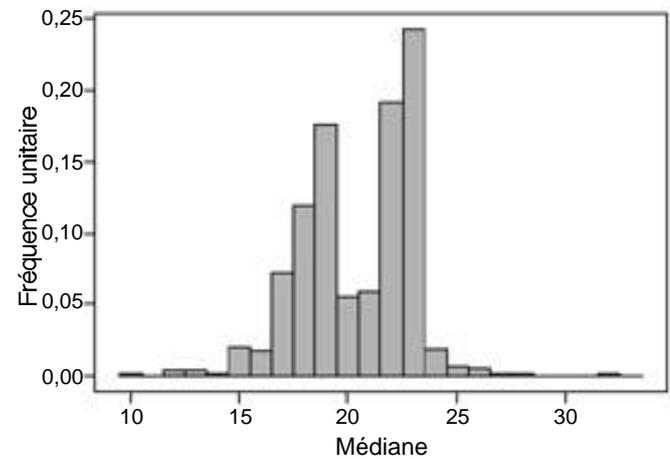


Figure 3. Distribution des médianes de 1.000 échantillons obtenus par *bootstrap* — *Distribution of the medians of 1,000 bootstrap samples.*

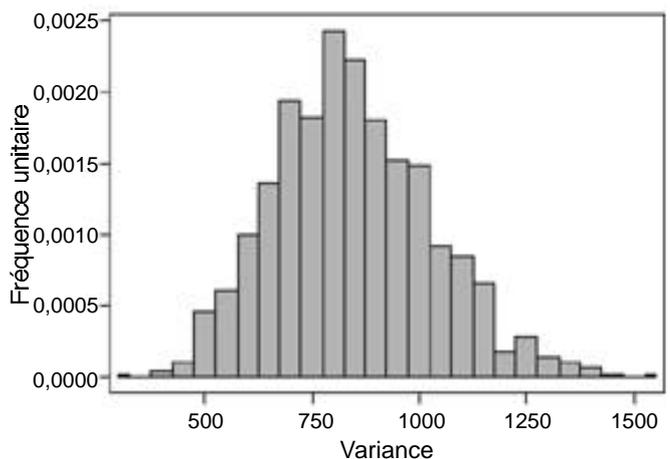


Figure 4. Distribution des variances de 1.000 échantillons obtenus par *bootstrap* — *Distribution of the variances of 1,000 bootstrap samples.*

l'écart-type des \hat{x}_k^* tend vers le résultat analytique, lorsque B tend vers l'infini.

Ainsi, pour la moyenne d'un échantillon aléatoire et simple, on sait que l'erreur-standard de la moyenne est égale à $\frac{\hat{\sigma}}{\sqrt{n}}$. Si B tend vers l'infini, l'écart-type $\hat{\sigma}_*$ tend vers $\hat{\sigma}_{plug}/\sqrt{n}$, avec :

$$\hat{\sigma}_{plug} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

L'estimation $\hat{\sigma}_{plug}$ de l'écart-type de la population donnée ci-dessus est appelée estimation par insertion (*plug-in estimator*). Pour un estimateur par insertion, la formule conduisant à l'estimation est la même que celle utilisée pour la définition du paramètre de la population. En effet, pour une population finie de taille N et de moyenne m_X , on a :

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - m_X)^2}$$

En d'autres mots, pour un estimateur par insertion, on considère l'échantillon comme une population particulière et on utilise la formule relative au paramètre de la population.

Pour l'exemple ci-dessus, l'écart-type des moyennes obtenues pour les différents échantillons x_k^* est égal à 2,89 (**Tableau 2**). Si on augmentait indéfiniment le nombre de répétitions B , cet écart-type se rapprocherait de 2,91 puisque :

$$\hat{\sigma}_{plug}^2 = (n-1) \hat{s}^2/n = (99)(854,63)/100 = 846,08$$

$$\text{et } \sqrt{\hat{\sigma}_{plug}^2/n} = \sqrt{846,08/100} = 2,91.$$

D'une manière générale, lorsque B tend vers l'infini, la valeur $\hat{\sigma}_*$ tend vers une valeur fixée qui correspond à l'estimation de l'erreur-standard du *bootstrap* idéal. Efron et Tibshirani (1993) proposent les règles empiriques suivantes pour le choix de B :

- un nombre réduit de répétitions ($B = 25$, par exemple) permet d'obtenir une première information et $B = 50$ est généralement suffisant pour avoir une bonne estimation de l'erreur-standard ;
- il est très rare que plus de 200 répétitions soient nécessaires pour estimer une erreur-standard.

On peut noter que le choix de B n'est pas fonction de la taille n de l'échantillon.

3.2. Estimation du biais

Le biais d'un paramètre peut être estimé par la méthode du *bootstrap* de la manière suivante :

$$\text{biais}_B(\hat{\theta}) = \hat{\theta} - \hat{\theta}_{plug}$$

$\hat{\theta}_{plug}$ étant l'estimation par *plug-in* du paramètre θ , même si $\hat{\theta}$ n'est pas une estimation par *plug-in*.

Ainsi, pour la moyenne et pour la médiane, $\hat{\theta}$ est égale à $\hat{\theta}_{plug}$ et l'estimation du biais pour ces deux paramètres est égale à :

$$27,99 - 28,13 = -0,14 \quad \text{et} \quad 20,44 - 21,56 = -1,12,$$

Pour la variance, l'estimation par insertion est égale à 846,08 et l'estimation du biais est égale à :

$$843,58 - 846,08 = -2,50.$$

Disposant d'une estimation du biais, on peut éventuellement corriger l'estimation initiale. On obtient alors :

$$\hat{\theta}_C = \hat{\theta} - \text{biais}_B(\hat{\theta}).$$

On notera cependant que la correction systématique du biais, par la relation ci-dessus, peut s'avérer dangereuse dans la mesure où cette correction peut augmenter l'erreur-standard de manière importante. En pratique, lorsque le rapport du biais à l'erreur-standard est inférieur à 0,25, il est souvent préférable de ne pas corriger l'estimateur pour le biais. D'autre part, un rapport supérieur à 0,25 peut être une indication que la statistique $\hat{\theta} = f(x_1, \dots, x_n)$ est inappropriée pour estimer θ .

Pour l'exemple ci-dessus, il n'y a certainement pas intérêt à corriger les estimations pour le biais, puisqu'on peut montrer que les trois estimations sont non biaisées (Dagnelie, 1998).

Lorsque $\hat{\theta}$ est une estimation par insertion, on a démontré qu'une meilleure estimation du biais est obtenue en tenant compte des proportions d'apparition P_i^* , définies au paragraphe 2.1. Il suffit de remplacer, dans la formule ci-dessus, $\hat{\theta}_{plug}$ par $\hat{\theta}_{plug}^*$, $\hat{\theta}_{plug}^*$ étant une fonction des observations de l'échantillon \mathbf{x} qui attribue aux observations x_i un poids égal à P_i^* , au lieu de leur attribuer des poids égaux à $P_0 = 1/n$.

3.3. Estimations par *jackknife*

Une autre forme de rééchantillonnage permettant d'estimer l'erreur-standard et le biais d'un paramètre est la technique du *jackknife*. On calcule n fois la valeur du paramètre à partir d'un échantillon de $n-1$ observations, chacune des observations étant éliminée à son tour de l'échantillon. En désignant par $\hat{\theta}_{(-i)}$ l'estimation obtenue après élimination de l'observation i , on peut estimer l'erreur-standard du paramètre θ par la relation suivante (Efron, Tibshirani, 1993 ; Dagnelie, 1998) :

$$\hat{\sigma}_J = \sqrt{\frac{n-1}{n} \sum_{i=1}^n (\hat{x}_{(-i)} - \hat{x}_J)^2}$$

avec :

$$\hat{x}_J = \frac{1}{n} \sum_{i=1}^n \hat{x}_{(-i)}$$

et le biais vaut :

$$\text{biais}_J(\hat{x}) = (n-1)(\hat{x}_J - \hat{x})$$

Lorsque la taille n de l'échantillon est inférieure au nombre de répétitions utilisé dans la méthode du *bootstrap*, soit le plus souvent entre 50 et 200, le calcul de l'erreur-standard et du biais est plus rapide par le *jackknife*. Par contre, la méthode du *jackknife*, qui peut être considérée comme une approximation du *bootstrap*, donne, de manière générale, de moins bonnes estimations.

De plus, la méthode du *jackknife* peut donner des résultats aberrants lorsque la statistique \hat{x} n'est pas une fonction continue (*smooth function*) des observations x_i , c'est-à-dire une fonction variant de manière régulière lorsque les données changent. La médiane est un cas typique de fonction non continue.

Pour l'exemple des exploitations agricoles, la méthode du *jackknife* donne les erreurs-standards suivantes : 2,92 pour la moyenne, 5,62 pour la médiane, 196,91 pour la variance.

Pour la moyenne et pour la variance, les erreurs-standards sont proches des valeurs obtenues par *bootstrap*, respectivement égales à 2,89 et 184,25 (Tableau 2). Par contre, pour la médiane, l'erreur-standard obtenue par *jackknife* est plus de deux fois plus grande que celle obtenue par *bootstrap*. Cette valeur est cependant aberrante, comme signalé ci-dessus, car elle n'est en fait fonction que de la différence qui existe entre deux observations successives particulières. Considérons, en effet, la série statistique constituée des six observations suivantes, ordonnées par ordre croissant :

18,91 ; 19,46 ; 21,00 ; 22,13 ; 22,16 et 22,65.

En éliminant tour à tour une observation, la médiane $\hat{x}_{(-i)}$ vaut soit 22,13 si l'observation éliminée est une des trois premières, soit 21,00 si l'observation éliminée est une des trois dernières. La somme des carrés des écarts de toutes ces médianes n'est donc fonction que de l'écart qui existe entre l'observation de rang $n/2$ et celle de rang $n/2 + 1$. La série statistique considérée ici est en réalité la partie centrale de l'échantillon x de 100 observations classées par ordre croissant. Plus précisément, il s'agit des observations de rang 48 à 53. En appliquant la méthode du *jackknife* à cet échantillon complet, on obtient 50 médianes

égales à 21,00 et 50 médianes égales à 22,13. L'erreur-standard de la médiane s'écrit donc :

$$\hat{\sigma}_J = \sqrt{\frac{(99)(100)}{100} [(22,13 - 21,00)/2]^2} = 5,62$$

et ne dépend que de l'écart entre l'observation de rang 50 et celle de rang 51.

4. INTERVALLE DE CONFIANCE

4.1. Méthode de l'erreur-standard

Une première solution consiste à définir l'intervalle de confiance par la méthode de l'erreur-standard (*standard bootstrap confidence interval*) :

$$\hat{x} \pm u_{1-\alpha/2} \hat{\sigma}_*$$

$u_{1-\alpha/2}$ étant le pourcentile $1-\alpha/2$ de la distribution normale réduite et $1-\alpha$ étant le degré de confiance retenu.

Pour que cette approche soit satisfaisante, il faut que la distribution d'échantillonnage du paramètre étudié soit approximativement normale, que l'estimateur soit non biaisé, et que $\hat{\sigma}_*$ soit une bonne estimation de l'erreur-standard de la distribution du paramètre.

Le fait que ces conditions soient remplies ou non dépend des circonstances. La condition de normalité peut être vérifiée à partir de la distribution des \hat{x}_k^* et il peut être utile éventuellement d'effectuer une transformation de manière à rendre la distribution plus proche de la normale. Le biais de l'estimateur peut être estimé, comme nous l'avons vu au paragraphe 3.2, mais sa prise en compte risque d'augmenter la variance de l'estimateur. Enfin, la qualité de l'estimation de l'erreur-standard est liée au nombre de répétitions B considéré et nous avons signalé, au paragraphe 3.1, que 50 répétitions sont généralement suffisantes.

Pour l'exemple considéré et pour un degré de confiance de 95 %, on obtient les limites de confiance suivantes, respectivement pour la moyenne, pour la médiane et pour la variance :

$$\begin{aligned} &28,13 \pm (1,96)(2,89), \text{ soit } 22,47 \text{ et } 33,79, \\ &21,56 \pm (1,96)(2,53), \text{ soit } 16,60 \text{ et } 26,52 \\ &\text{et } 854,63 \pm (1,96)(184,25), \text{ soit } 493,50 \text{ et } 1215,76. \end{aligned}$$

Le caractère tout à fait non normal de la distribution des médianes nous conduit cependant à mettre en doute, *a priori*, le résultat obtenu pour ce paramètre.

4.2. Méthode des pourcentiles simples

Dans la méthode des pourcentiles simples (*simple percentile confidence interval*), les limites de confiance

sont données par les pourcentiles $1/2$ et $1 - 1/2$ de la distribution d'échantillonnage empirique $\hat{\lambda}_k^*$, c'est-à-dire de la distribution des $\hat{\lambda}_k^*$. Nous les notons $\hat{\lambda}_{[1/2]}^*$ et $\hat{\lambda}_{[1-1/2]}^*$.

Contrairement à la méthode de l'erreur-standard, la distribution d'échantillonnage du paramètre étudié ne doit pas être normale pour que la méthode des pourcentiles soit satisfaisante. Par contre, le nombre de rééchantillonnages B doit être plus élevé que dans le cas de la méthode de l'erreur-standard, car il faut un plus grand nombre d'observations pour estimer, avec une précision suffisante, un pourcentile que pour estimer un écart-type. B sera par exemple de l'ordre de 1.000.

Pour 1.000 rééchantillonnages et pour un degré de confiance de 95 %, les pourcentiles 0,025 et 0,975 correspondent approximativement à l'observation de rang 25 et à l'observation de rang 975, la valeur exacte pouvant dépendre de l'algorithme utilisé pour le calcul de ces pourcentiles. Les résultats obtenus pour les trois paramètres considérés dans l'exemple sont les suivants : 22,48 et 33,89 pour la moyenne, 14,82 et 24,01 pour la médiane, et 512,92 et 1248,80 pour la variance.

Il faut noter aussi qu'une procédure de calcul un peu différente a été proposée par Hall (1992) et est décrite par Manly (1997). La méthode consiste à calculer les écarts :

$$\hat{e}_k^* = \hat{\lambda}_k - \hat{\lambda}_k^*$$

et à déterminer les pourcentiles $1/2$ et $1 - 1/2$, notés $\hat{e}_{[1/2]}^*$ et $\hat{e}_{[1-1/2]}^*$ de cette distribution. Les limites de confiance sont alors données par les relations :

$$\hat{\lambda} - \hat{e}_{[1-1/2]}^* \quad \text{et} \quad \hat{\lambda} + \hat{e}_{[1/2]}^*$$

Par cette méthode on obtient, dans le cas de l'exemple 22,37 et 33,78 pour la moyenne, 19,11 et 28,31 pour la médiane, et 460,46 et 1196,34 pour la variance.

4.3. Méthode des pourcentiles corrigés pour le biais

On détermine d'abord la proportion p de valeurs $\hat{\lambda}_k^*$ inférieures à $\hat{\lambda}$ et on calcule le pourcentile u_p relatif à la distribution normale réduite.

Soit u_1 et u_2 les valeurs de la fonction de répartition de la normale réduite aux points u_1 et u_2 :

$$u_1 = \Phi(u_1) \quad \text{et} \quad u_2 = \Phi(u_2),$$

avec $u_1 = 2u_p + u_{1/2}$ et $u_2 = 2u_p + u_{1-1/2}$.

Les limites de confiance déterminées par la méthode des pourcentiles corrigés pour le biais (*bias corrected percentile confidence interval*) sont alors les pourcentiles $\hat{\lambda}_{[u_1]}^*$ et $\hat{\lambda}_{[u_2]}^*$ de la distribution des $\hat{\lambda}_k^*$. Des informations concernant l'origine de cette

correction sont données dans Efron et Tibshirani (1993) et dans Chernick (1999).

On remarque que si $p = 0,5$, c'est-à-dire si $\hat{\lambda}$ est la médiane de la distribution des $\hat{\lambda}_k^*$, il n'y a pas de correction pour le biais, puisque $u_p = 0$, et on retrouve la méthode précédente. Si p est inférieur à 0,5, les limites de confiance correspondent à des pourcentiles inférieurs respectivement à $1/2$ et $1 - 1/2$. Au contraire, si p est supérieur à 0,5, les limites correspondent à des pourcentiles supérieurs à $1/2$ et $1 - 1/2$. Par exemple, si $p = 0,4$, les limites de confiance, pour un degré de confiance de 0,95, sont les pourcentiles $\hat{\lambda}_{[0,0068]}^*$ et $\hat{\lambda}_{[0,9269]}^*$. Pour $p = 0,6$, les limites correspondent à $\hat{\lambda}_{[0,0731]}^*$ et $\hat{\lambda}_{[0,9932]}^*$.

Pour les exploitations agricoles, les valeurs de p sont égales à 0,542, 0,533 et 0,555 respectivement pour la moyenne, pour la médiane et pour la variance. Les limites de confiance seront par conséquent toutes légèrement plus grandes que dans le cas de la méthode des pourcentiles simples. On obtient en effet 23,03 et 34,62 pour la moyenne, 15,78 et 24,17 pour la médiane et 554,15 et 1303,12 pour la variance.

4.4. Méthode des pourcentiles avec correction pour le biais et accélération

La méthode précédente, qui prend en compte le biais, peut être étendue de manière à tenir compte d'un éventuel changement de l'erreur-standard de $\hat{\lambda}$ lorsque $\hat{\lambda}$ varie. Elle porte alors le nom de méthode des pourcentiles avec correction pour le biais et accélération (*bias corrected and accelerated confidence interval*). Une justification de cette méthode est donnée par Efron et Tibshirani (1993).

Les limites de confiance sont les pourcentiles $\hat{\lambda}_{[u_1]}^*$ et $\hat{\lambda}_{[u_2]}^*$ de la distribution des $\hat{\lambda}_k^*$, u_1 et u_2 étant les valeurs de la fonction de répartition de la variable normale réduite aux points u_1 et u_2 définis de la manière suivante :

$$u_1 = u_p + (u_p + u_{1/2}) / [1 - a(u_p + u_{1/2})]$$

$$u_2 = u_p + (u_p + u_{1-1/2}) / [1 - a(u_p + u_{1-1/2})]$$

Dans ces relations, u_p est défini comme précédemment et la constante a est appelée accélération, car elle est liée au taux de variation de l'erreur-standard de $\hat{\lambda}$ lorsque le paramètre $\hat{\lambda}$ varie. Cette constante peut être estimée de différentes manières. Une solution consiste à utiliser la technique du *jackknife*. On obtient alors le paramètre a par la relation suivante :

$$a = \frac{\sum_{i=1}^n (\hat{\lambda}_J - \hat{\lambda}_{(-i)})^3}{6 \sum_{i=1}^n (\hat{\lambda}_J - \hat{\lambda}_{(-i)})^2} \quad 3/2$$

Dans cette relation, $\hat{\mu}_{(-i)}$ est l'estimation du paramètre obtenue à partir de l'échantillon initial, dont on a enlevé la $i^{\text{ème}}$ observation et $\hat{\mu}_J$ est la moyenne des n valeurs $\hat{\mu}_{(-i)}$.

On peut constater que si $a = 0$, on retrouve la méthode des pourcentiles corrigés pour le biais. La prise en compte de l'accélération constitue donc bien une extension de la méthode précédente.

Les limites obtenues par cette méthode sont, pour l'exemple examiné, 23,21 et 35,33 pour la moyenne, 15,78 et 24,17 pour la médiane et 589,75 et 1405,85 pour la variance.

On constate que pour la médiane les résultats sont identiques à ceux obtenus par la méthode précédente, le paramètre a étant nul, du fait de la parfaite symétrie de la distribution des $\tilde{x}_{(-i)}$, comme nous l'avons précisé au paragraphe 3.3.

4.5. Méthode du bootstrap-t

L'idée mise en application dans la *bootstrap-t* est de définir une statistique dont la distribution ne soit pas fonction de la valeur réelle et inconnue du paramètre μ . La statistique T :

$$T = \frac{\hat{\mu} - \mu}{\hat{\sigma}_k}$$

peut remplir ce rôle.

Il s'agit alors d'approcher la distribution théorique de T par rééchantillonnage. Dans ce but, on calcule :

$$t_k = \frac{\hat{\mu}_k - \hat{\mu}}{\hat{\sigma}_k}$$

$\hat{\sigma}_k$ ($\hat{\sigma}_k^*$) étant l'erreur-standard de $\hat{\mu}_k$ ($\hat{\mu}_k^*$), qui dépend donc de l'échantillon k . Elle peut être calculée par une formule théorique, lorsqu'une telle formule est disponible, ou à partir du rééchantillonnage de l'échantillon x_k^* utilisé pour calculer $\hat{\mu}_k$. Il s'agit alors d'un *bootstrap* à deux niveaux, puisque chaque échantillon x_k^* fait lui-même l'objet d'un rééchantillonnage, permettant de calculer l'erreur-standard selon la méthode décrite au paragraphe 3.1.

Une autre possibilité pour estimer l'erreur-standard est le recours à la méthode du *jackknife* (paragraphe 3.3). Pour l'échantillon x_k^* , on détermine les n valeurs $\hat{\mu}_{(-i)}^*$, en éliminant à tour de rôle chacune des observations, et on calcule l'erreur-standard du paramètre par la relation suivante :

$$\hat{\sigma}_k^* = \sqrt{\frac{n-1}{n} \sum_{i=1}^n (\hat{\mu}_{(-i)}^* - \hat{\mu}_k^*)^2}$$

Lorsque l'effectif n de l'échantillon est inférieur au nombre de répétitions au second niveau, noté B , la

méthode du *jackknife* exige moins de calculs, mais l'estimation est généralement moins bonne que l'estimation par *bootstrap*, du moins lorsque B est suffisamment grand (de 50 à 200, par exemple), comme signalé au paragraphe 3.3.

Disposant de la distribution des t_k^* , on en détermine les pourcentiles $1/2$ et $1 - 1/2$ notés $t_{[1/2]}^*$ et $t_{[1-1/2]}^*$, et on obtient les limites de confiance du paramètre μ par les relations :

$$\hat{\mu}_1 = \hat{\mu} - t_{[1-1/2]}^* \hat{\sigma}_k \quad \text{et} \quad \hat{\mu}_2 = \hat{\mu} - t_{[1/2]}^* \hat{\sigma}_k$$

Le *bootstrap* à deux niveaux a été appliqué à l'échantillon des 100 tailles d'exploitations agricoles, le nombre de répétitions au deuxième niveau étant égal à 50. Pour la moyenne, les pourcentiles 0,025 et 0,975 de la distribution observée des t_k^* sont les suivants :

$$t_{[0,025]}^* = -2,3364 \quad \text{et} \quad t_{[0,975]}^* = 1,7574$$

Les limites de confiance de la moyenne sont par conséquent égales à :

$$28,13 - (1,7574)(2,89) = 23,05$$

$$\text{et } 28,13 - (-2,3364)(2,89) = 34,88$$

Des calculs identiques ont été réalisés pour la médiane et pour la variance ; les limites de confiance obtenues sont respectivement 18,45 et 27,66 pour la médiane et 561,19 et 1523,14 pour la variance.

4.6. Choix d'une méthode

Dans les paragraphes précédents, diverses méthodes de calcul de l'intervalle de confiance ont été décrites. Le **tableau 3** reprend quelques caractéristiques de ces méthodes. La méthode du pourcentile avec correction

Tableau 3. Caractéristiques des méthodes de calcul de l'intervalle de confiance d'un paramètre (1 : méthode de l'erreur-standard ; 2 : méthode des pourcentiles simples ; 3 : méthode des pourcentiles corrigés pour le biais et accélération ; 4 : méthode du *bootstrap-t*) — *Characteristics of the methods for computing confidence interval of a parameter* (1: standard bootstrap; 2: simple percentile; 3: bias corrected and accelerated; 4: bootstrap-t confidence interval).

Méthodes	1	2	3	4
Nombre de rééchantillonnages	100	1.000	1.000	1.000 × 100
Influence d'une transformation	oui	non	non	oui
Respect du domaine	non	oui	oui	non
Ordre de précision	$n^{-1/2}$	$n^{-1/2}$	n^{-1}	n^{-1}

du biais n'a pas été reprise car il s'agit d'un cas particulier de la méthode avec correction du biais et accélération.

La première ligne du tableau concerne l'ordre de grandeur du nombre d'échantillons qu'il faut prélever pour le calcul des limites de confiance. La méthode de l'erreur-standard est la plus rapide ($B = 100$, par exemple) alors que la méthode du *bootstrap-t* est la plus coûteuse, puisqu'elle fait appel au *bootstrap* à deux niveaux : par exemple, $B = 1.000$ répétitions au premier niveau, chacune de celles-ci faisant l'objet de $B' = 100$ répétitions pour l'estimation de l'erreur-standard, sauf si n est plus petit que 100 et qu'on utilise la méthode du *jackknife*, qui est cependant en général moins bonne.

La deuxième ligne du tableau concerne l'influence d'une transformation, comme par exemple la transformation argument tangente hyperbolique dans le cas du coefficient de corrélation. La méthode de l'erreur-standard et la méthode du *bootstrap-t* sont influencées par une transformation alors que les méthodes basées sur les pourcentiles ne le sont pas. Aucune transformation ne doit donc être recherchée pour les dernières méthodes citées.

La troisième ligne du tableau signale si l'intervalle de confiance respecte le domaine dans lequel doit se trouver le paramètre. Ainsi, par exemple, les méthodes qui ne respectent pas le domaine (méthode de l'erreur-standard et méthode du *bootstrap-t*) peuvent conduire à des limites de confiance qui ne sont pas dans le domaine $(-1, 1)$ pour un coefficient de corrélation, alors que les autres méthodes ne donneront jamais des limites de confiance situées en dehors de ce domaine.

Enfin, la quatrième ligne a trait à la proportion d'intervalles de confiance corrects. Idéalement, pour un degré de confiance égal à 0,95 par exemple, la probabilité que la limite inférieure de l'intervalle de confiance soit supérieure à doit être égale à 0,025. De même, la probabilité que la limite supérieure de l'intervalle soit inférieure à doit être égale à 0,025.

Plutôt que de considérer les deux limites simultanément, on peut s'intéresser à une seule limite. Soit $[]$ la limite calculée, telle que, idéalement :

$$P(< []) = \dots$$

On peut montrer que les méthodes de l'erreur-standard et des pourcentiles simples conduisent à des probabilités, en général, égales à :

$$P(< []) = \dots + (n^{-1/2}),$$

tandis que les méthodes du *bootstrap-t* et des pourcentiles avec correction du biais et accélération sont du type :

$$P(< []) = \dots + (n^{-1}).$$

Les termes $(n^{-1/2})$ et (n^{-1}) indiquent respectivement l'ordre de grandeur de l'erreur. Ces relations montrent que les deux dernières méthodes citées sont préférables de ce point de vue.

La comparaison des différentes méthodes proposées montre que la méthode des pourcentiles corrigés pour le biais et accélération offre, dans l'ensemble, le plus d'avantages et est, de ce fait, préconisée par Efron et Tibshirani [1993].

4.7. Comparaison des résultats pour l'exemple

Les limites de confiance obtenues par les différentes méthodes ont été données dans les paragraphes précédents (paragraphes 4.1 à 4.5).

À titre de comparaison, les intervalles de confiance ont également été calculés en utilisant des méthodes paramétriques classiques, même si, *a priori*, on sait que les conditions d'application ne sont pas remplies. Ainsi, pour la moyenne on a :

$$\bar{x} \pm t_{1-\alpha/2} \hat{\sigma} / \sqrt{n}.$$

Pour la médiane, et pour une population-parent normale, on a :

$$\tilde{x} \pm u_{1-\alpha/2} \sqrt{\hat{\sigma}^2 / 2n}.$$

Enfin, pour la variance, on a retenu la méthode basée sur l'utilisation de la distribution χ^2 à $n - 1$ degrés de liberté, valable pour une population normale, et la méthode de l'erreur-standard :

$$\hat{\sigma}^2 \pm u_{1-\alpha/2} \sqrt{V(\hat{\sigma}^2)},$$

avec

$$V(\hat{\sigma}^2) = [(n - 1) \hat{\sigma}^2 - n + 3] \hat{\sigma}^4 / (n - 1)n,$$

qui suppose que la distribution d'échantillonnage de la variance est proche de la distribution normale.

Le **tableau 4** reprend les limites de confiance pour les trois paramètres envisagés. On ne note pas de différences importantes entre les méthodes pour la moyenne. Par contre, les différences sont plus importantes pour la médiane. En particulier, l'intervalle obtenu par la méthode classique est sensiblement plus large que celui obtenu par les méthodes basées sur le rééchantillonnage. Pour la variance, les résultats sont également assez différents selon la méthode utilisée : la méthode classique basée sur l'utilisation de la distribution χ^2 présente un intervalle nettement plus étroit et la méthode des

Tableau 4. Intervalles de confiance de la moyenne, de la médiane et de la variance : résultats pour les différentes méthodes — *Confidence intervals of the mean, the median and the variance: results for the various methods.*

Méthodes	Moyenne	Médiane	Variance
Bootstrap			
Erreur-standard	22,47 – 33,79	16,60 – 26,52	494 – 1.216
Pourcentiles simples	22,48 – 33,89	14,82 – 24,01	513 – 1.249
Correction/biais	23,03 – 34,62	15,78 – 24,17	554 – 1.303
Correction/ biais et accélération	23,21 – 35,33	15,78 – 24,17	590 – 1.406
<i>Bootstrap-t</i>	23,05 – 34,88	18,45 – 27,66	561 – 1.523
Méthodes classiques			
Erreur-standard	22,33 – 33,93	14,39 – 28,74	465 – 1.244
Distribution ²	-	-	659 – 1.153

pourcentiles avec correction du biais et accélération et surtout le *bootstrap-t* conduisent à un intervalle plus large que les autres méthodes.

Pour mieux apprécier la qualité des différentes méthodes comparées, Sodjinou (2001) a répété les calculs pour 1.000 échantillons aléatoires et simples prélevés dans la population-parent et a déterminé la proportion d'intervalles de confiance ne contenant pas la valeur réelle du paramètre. Les résultats obtenus sont repris dans le **tableau 5**. Les intervalles étant calculés pour un niveau de confiance théorique de 0,95, la proportion d'intervalles ne contenant pas le paramètre réel devrait idéalement être de $\alpha = 0,05$.

Pour la moyenne, on note une bonne concordance entre les différentes méthodes, la proportion d'intervalles ne contenant pas la moyenne étant

Tableau 5. Pourcentages d'intervalles de confiance ne contenant pas le paramètre de la population : résultats pour les différentes méthodes — *Percentages of confidence intervals that do not include the parameter of the population: results for various methods.*

Méthodes	Moyenne	Médiane	Variance
Bootstrap			
Erreur-standard	6,3	5,6	22,9
Pourcentiles simples	6,1	5,1	21,7
Correction/biais	6,0	5,1	19,9
Correction/biais et accélération	6,7	5,2	15,8
<i>Bootstrap-t</i>	5,2	6,1	9,8
Méthodes classiques			
Erreur-standard	5,8	4,5	22,0
Distribution ²	-	-	44,5

légèrement trop grande pour l'ensemble des méthodes.

Pour la médiane, on note également une bonne concordance entre les résultats. La méthode classique de l'erreur-standard conduit à des résultats comparables à ceux obtenus par *bootstrap* bien que la formule d'estimation de la variance de la médiane qui a été utilisée ne soit valable que pour des populations normales. Signalons cependant que la méthode de l'erreur-standard n'est pas la seule méthode paramétrique disponible. Ainsi, Minitab (2000) propose une solution basée sur l'utilisation de la distribution binomiale de paramètre $p = 0,5$.

Dans le cas de la variance, on constate l'inadéquation de la méthode basée sur la variable χ^2 , pour laquelle 45 % des intervalles ne contiennent pas la variance de la population. La méthode est donc très mauvaise dans le cas d'une non-normalité accentuée de la population-parent. La méthode classique de l'erreur-standard est un peu moins mauvaise et donne des résultats comparables aux méthodes basées sur le rééchantillonnage, à l'exception du *bootstrap-t*, qui pour cet exemple est sensiblement supérieur aux autres, bien que la proportion d'intervalles ne contenant pas la moyenne soit le double de la valeur attendue.

5. CONCLUSIONS

“Le *bootstrap* est une méthode d'inférence statistique basée sur l'utilisation de l'ordinateur qui peut répondre sans formules à beaucoup de questions statistiques réelles”. C'est en ces termes qu'Efron et Tibshirani (1993) présentent le *bootstrap* dans la préface de leur ouvrage.

Il est incontestable que l'utilisation des techniques de rééchantillonnage a été rendue possible grâce à la généralisation des moyens de calculs performants. Ces techniques reposent, au départ, sur des idées simples. Toutefois, il faut bien admettre que les développements apportés aux méthodes de base leur ont fait perdre une partie de cette simplicité.

Dans cette note, nous nous sommes limité au problème de l'estimation du biais et de l'erreur-standard d'un paramètre, et à la détermination des limites de confiance d'un paramètre. Il ne s'agit cependant pas là des seules applications des méthodes de rééchantillonnage. Celles-ci peuvent, en effet, aussi être utilisées pour la réalisation de différents tests d'hypothèses, pour le choix des variables et l'estimation de l'erreur de prédiction en régression, pour l'estimation du taux d'erreur en analyse discriminante, notamment.

Bien qu'elles puissent être utilisées dans des situations très variées, leur mise en œuvre ne présente guère d'intérêt lorsque l'inférence statistique peut être réalisée par des méthodes analytiques classiques, pour lesquelles les conditions d'application sont remplies. Elles ne sont donc pas destinées à remplacer les

méthodes d'inférence statistique classiques lorsque celles-ci sont applicables mais plutôt à fournir des réponses à des questions pour lesquelles les méthodes classiques sont inapplicables ou non disponibles. Ainsi, pour l'exemple traité, le recours au *bootstrap* ne se justifie certainement pas dans le cas de la moyenne et de la médiane. Pour la moyenne, on peut en effet utiliser la méthode de l'erreur-standard, compte tenu du caractère approximativement normal de la distribution d'échantillonnage de la moyenne, vu la taille de l'échantillon. Pour la médiane, on dispose d'une approche non paramétrique (Minitab, 2000). Par contre, pour la variance, l'utilisation du *bootstrap* constitue une alternative valable.

Le caractère relativement général des problèmes qui peuvent être résolus par *bootstrap* ne doit pas faire perdre de vue que la qualité de l'inférence dépend de la nature de la question posée et de la disponibilité des données. Comme le suggère Manly (1997), le *bootstrap* doit être utilisé avec prudence dans les situations où il n'a pas encore été testé de manière approfondie.

Une discussion assez générale sur l'intérêt et les limites du *bootstrap* est donnée dans la synthèse proposée par Young (1994) et dans les commentaires suscités par l'article en question.

Enfin, l'absence dans les logiciels classiques de fonctions permettant d'utiliser le *bootstrap*, sans obligation pour l'utilisateur de programmer lui-même les procédures de calcul, est certainement un frein à l'utilisation du *bootstrap* dans les cas où il pourrait être utilisé.

Bibliographie

- Chernick MR. (1999). *Bootstrap methods: a practitioner's guide*. New York: Wiley, 264 p.
- Dagnelie P. (1998). *Statistique théorique et appliquée. Tome 1 : statistique descriptive et bases de l'inférence statistique*. Bruxelles : De Boeck et Larcier, 508 p.
- Efron B., Tibshirani RJ. (1993). *An introduction to the bootstrap*. New York: Chapman and Hall, 436 p.
- Hall P. (1992). *The bootstrap and Edgeworth expansion*. New York: Springer, 352 p.
- INS (1996). *Recensement agricole et horticole au 15 mai 1995. Statistiques Agricoles*. Bruxelles : Institut National de Statistique, p. 196.
- Léger C., Politis DHN., Romano JP. (1992). Bootstrap technology and applications. *Technometrics* **34**, p. 378–398.
- Manly BFJ. (1997). *Randomization, bootstrap and Monte Carlo methods in biology*. New York: Chapman and Hall, 399 p.
- Minitab (2000). *Minitab user's guide 2: data analysis and quality tools. Release 13 for Windows*. Minitab.
- Sodjinou (2001). *Techniques de rééchantillonnage avec le logiciel statistique S-Plus 2000 : application à la régression*. Travail de fin d'études. Fac. univ. Sci. agron. Gembloux, 113 p.
- Young GA. (1994). Bootstrap: more than a stab in the dark? *Stat. Sci.* **9**, p. 382–415.

(10 réf.)