الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique Et POPULAIRE
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieure et de la Recherche Scientifique
جامعة سعد دحلب البليدة 1
Université Saad Dahleb de Blida 1
Faculté des Sciences
Département des Mathématiques

Mémoire de Fin d'Etudes

En vue de l'obtention du diplôme de MASTER

En : Mathématiques

Option : MODELISATION STOCHASTIQUE ET STATISTIQUE

Présentée par :

❖ Hamamti Sabrine
❖ Sidaoui Imene

THEME

Echantillonnage par ensembles classés

Soutenu le 24 /07/2019,

Devant les jury composé de :

Mr O.TAMI	MAA	**USD Blida 1**	**Président**
Mr A. RASSOUL	MCA	**ENSH Blida 1**	**Promoteur**
Mr R. FRIHI	MAA	**USD Blida 1**	**Examinateur**

Année universitaire : 2018/2019.

Table des matières

Table des figures

Liste des tableaux

0.1 Dédicace

Je dédie ce modeste travail : Aux deux êtres humains qui sont les plus chères dans ma vie
À ma lumière, celle qui m'a donnée la vie, l'amour, la tendresse et le courage, toi chère mère
Et celui qui m'a soutenu et guidé afin que je puisse arriver à cette étape de ma vie, toi chère père
Je dédie aussi cette modeste réalisation à :
Mes trés chères soeurs. Je n'oublie pas mes tantes ; mes oncles et à toutes la familles : Hamamti ; Djilli.
A mon encadreur RASSOUL Abdelaziz, en espérant qu'il trouvre dans ce travail le témoignage de ma profonde gratitude.
A tout mes enseignants de l'université de Saad Dahleb sans exception.
A tous mes collègues et mes chers amis.
A tout ceux qui m'aiment, m'aident, m'encouragent toujours pour continuer sur la bonne voie .
H.Sabrine
Je dédie ce modeste travail
A ma maman qui m'a soutenu et encoragé durant ces années d'études.
A mon père qui m'a donnée *le courage.*
Qu'elle trouve ici le témoignage de ma profonde reconnaissance.
A mon frère mes soeurs , mes grands parent et ceux qui ont partage avec moi tous les moments d'émotion lors de la réalisation de ce travail . Ils m'ont chaleureusement supporté et encouragé tout au long de mon parcours.
A mon encadreur RASSOUL Abdelaziz, en espérant qu'il trouvre dans ce travail le témoignage de ma profonde gratitude.
A tout mes enseignants de l'université de Saad Dahleb sans exception.
A toutes la familles :Sidaoui ;Bouazri, mes proches et à ceux qui me donnent de l'amour et de la vivacité.
Atous mes amis qui m'ont toujours encouragé, et qui je souhaite plus de succés
A tous ceux que j'aime.
S.Imene

0.2 Remerciemment

Je remercie ALLAH de nous avoir donné la force et le courage de mener à bout ce travail.

Je tiens à exprimer toute ma reconnaissance à mon encadreur RASSOUL Abdelaziz. Je le remercie de m'avoir encadré, orienté, aidé et conseillé.

J'adresse mes sincères remerciements à tous les professeurs, intervenants et toutes les personnes qui par leurs paroles, leurs écrits, leurs conseils et leurs critiques ont guidé mes réflexions et ont accepté à me rencontrer et répondre à mes questions durant mes recherches.

Je remercie mes très chers parents, qui ont toujours été là pour moi, « Vous avez tout sacrifié pour vos enfants n'épargnant ni santé ni efforts. Vous m'avez donné un magnifique modèle de labeur et de persévérance. Je suis redevable d'une éducation dont je suis fier ».

Je remercie mes sœurs pour leur encouragement.

Enfin, je remercie tous mes Ami(e)s que j'aime.

À tous ces intervenants, je présente mes remerciements, mon respect et ma gratitude.

ملخص
الهدف من هذا العمل هو تقديم طريقة مبتكرة لأخذ العينات العشوائية تعرف بالعينات المرتبة كبديل عن العينات العشوائية البسيطة.
نقوم بدراسة الخصائص الإحصائية كالأمل، الانحراف المعياري، الكثافة، دالة التوزيع و المكممات.
في الأخير نقدم دراسة عن طريق المحاكاة لتقدير هذه القيم الإحصائية.
كما نقوم بحساب مؤشرات الخطر بالنسبة لمؤشر البورصة CAC40 والمقارنة ب التقديرات المعروفة سلفا.

0.3 Résumé

L'objet de ce travail est de donner une étude sur la méthode d'échantillonnage appelée "l'échantillonnage classé," ou "ranked set sampling" en anglais. La technique est d'abord présentée dans le cas d'une population univariée dont on cherche à estimer l'espérance et la variance. On montre que sous des hypothèses assez faibles, la moyenne expérimentale d'un échantillon classé est sans biais et que sa variance est inférieure à celle de l'estimateur traditionnel basé sur un échantillonnage aléatoire simple de même taille. On montre que dans les mêmes conditions que pour le cas univarié, l'efficacité de l'estimateur fondé sur l'échantrillon bivarié est supérieure à celle de la moyenne bivariée calculée à partir d'un échantillonnage aléatoire simple.

0.4 Abstract

The purpose of this work is to give a study on the sampling method called "ranked set sampling" The technique is first presented in the case of a univariate population whose hope and variance are estimated. It is shown that under fairly low assumptions, the experimental mean of a ranked sample is unbiased and its variance is lower than that of the traditional estimator based on simple random sampling of the same size. It is shown that under the same conditions as for the univariate case, the efficiency of the estimator based on the bivariate sample is greater than that of the bivariate average calculated from a simple random sampling.

0.5 Introduction Générale

Au début des années 50, McIntyre proposé une méthode d'échantillonnage qui, ultérieurement, est devenu connu sous le nom d'échantillonnage classés (Ranked Set Sampling RSS). La notion de RSS fournit un moyen efficace de réaliser une économie d'observation dans certains cas particuliers. Bien que la méthode soit restée en sommeil pendant longtemps, sa valeur a été redécouverte au cours des 20 dernières années environ en raison de son caractère rentable.Il y a eu beaucoup de nouveaux développements de l'idée originale de McIntyre,qui a rendu la méthode applicable dans un éventail de domaines beaucoup plus large que initialement prévu.La base théorique de RSS a été posée. Beaucoup des variantes et des ramifications de l'idée originale ont été proposées et étudiées.De plus en plus d'applications de RSS ont été citées dans la littérature.Le but de cette monographie est de donner un compte rendu systématique de la théorie et application de RSS. Dans cette introduction, nous donnons une brève discussion sur le notion de RSS et son applicabilité, une note historique sur le développement de RSS depuis qu'il a été proposé par McIntyre et, enfin, un aperçu du contenu de la monographie.

Dans ce mémoire, qui s'articule autour de trois chapitres, nous avons étudié la méthode d'échantillonnage classé et leurs estimations .

Dans le premier chapitre, nous effectuons une synthèse sur l'échantillonnage classé.Nous donnons quelques définition sur l'échantillonnage ainsi les différentes méthodes et l'évaluation d'échantillonnages. Nous avons traité de l'échantrillonnage univarié. De plus, nous avons pu vérifier que l'estimateur de la moyenne échantillonnale classé est plus précis que celui de la moyenne échantillonnale.

Dans le deuxième chapitre, nous donnons une généralité et l'estimateur de quantile avec l'échantillonnage classé, puis nous étudions les procédures d'inférence pour les quantiles d'une population basées sur un échantillon classé.Nous avons étudier l'échantillonnage classé bivarié. Puis nous donnons les différentes mesures de risque.

Le troisième chapitre, est consacré à l'applications et simulation par la méthode RSS on a fait quelques simulations pour voir le comportement des estimateurs proposés avec logiciel R.Nous donnons quelques rappels sur les propriétés et les caractérisations d'une mesure de risque. Nous détaillons une mesure particulière : la Value-at-Risk en donnant ses différentes variantes et les liens entre elles. Nous avons donnée une application sur des données réel (CAC 40), et en fin nous donnons une conclusion générale.

Chapitre 1

Généralités sur l'échantillonage classé

1.1 Introduction

L'échantillonnage joue un rôle central en statistique et les méthodes d'échantillonnage se sont beaucoup développées au fil du temps. Parmi celles-ci, la plus connue et la plus simple est sans aucun doute l'échantillonnage aléatoire simple. Dans un échantillon de ce genre, toutes les unités de la population ont la méme probabilité d'étre choisies.

Bien que l'augmentation de la taille d'un échantillon aléatoire simple permet généralement d'accroitre la précision de l'inférence, ce mode de collecte de données n'est pas toujours le plus approprié, particulièrement lorsque l'on s'intéresse à des sous-populations.

Lorsqu'une analyse plus fine de sous-populations est envisagée, on fait typiquement appel à des méthodes d'échantillonnage plus structurées que l'échantillonnage cette méthode est appelée échantillonage classé ou bien { Ranked Set Sampling (RSS) } par McIntyre (1952) [13], cette stratégie d'échantillonnage consiste à recueillir, selon un certain protocole, un nombre d'observations beaucoup plus grand que celles sur lesquelles des mesures seront éventuellement prises.

Ces observations, regroupées en échantillons aléatoires simples, sont alors triées au sein de chaque groupe, mais sans avoir recours à un instrument de mesure. En supposant que cette opération puisse ètre effectuée correctement (et à faible cout), on est alors en mesure d'extraire de chaque groupe une statistique d'ordre, sur laquelle la ou les caractéristique(s) d'intérèt sera (seront) mesurée(s).

Le protocole d'échantrillonnage y sera précisé et quelques travaux clas-

siques portant sur cette technique seront relatés. En particulier, une comparaison y sera faite de l'fficacité d'une estimation de la moyenne d'une population, selon que les données ont été recueillies par échantillonnage aléatoire simple ou par échantrillonnage.

Et en fin L'objectif de l'échantillonnage est de prélever une partie représentative d'un ensemble ou d'un lot de minerai pour déterminer avec la plus grande précision possible la teneur moyenne en divers éléments de cet ensemble.

1.2 Définition d'un échantillonnage par ensemble classé

Le principe de base de RSS (ranked set sampling) est une population infinie étudiée et l'hypothèse selon laquelle un ensemble d'unités d'échantillonnage issues de la population peut être classé par certains moyens plutôt à moindre coût sans la mesure réelle de la variable d'intérêt, qui est coûteuse et prend du temps.

Cette hypothèse peut sembler plutôt restrictive à première vue, mais il s'avère que dans la pratique, il existe de nombreuses situations dans lesquelles cela est satisfait.

Définition 1.1 (Population) *Une population est un ensemble fini d'objets (unités statistique) sur lesquels une étude se porte et dont les éléments répondent à une ou plusieurs caractéristique communes .*

Définition 1.2 (Échantillon) *Un échantillon est un sous-ensemble d'une population, ou bien une partie d'un ensemble choisi pour représenter une ou plusieurs propriétés caractéristiques de cet ensemble .*

1.2.1 Comment procède -t-on à la séléction d 'un échantillon de taille k

Pour créer des ensembles classés, nous devons partitionner les échantillon sélectionné de première phase en ensembles de taille égale.Il faut donc choisir une taille d'ensemble généralement petite, environ 3 ou 4, pour minimiser les erreurs de classement. Appelons arbitrairement cette taille définie k, où k est le nombre d'unités d'échantillonnage allouées dans chaque ensemble. Maintenant procéder comme suit :

1 Sélectionnez au hasard k^2 unités d'échantillonnage de la population.

2 Répartissez les k^2 unités sélectionnées de manière aussi aléatoire que possible en k ensembles de taille k.

3 Sans connaître encore aucune valeur pour la variable d'intérêt, classez les unités de chaque ensemble par rapport à la variable d'intérêt. Cela peut être basé sur jugement professionnel personnel ou avec une variable concomitante en corrélation avec la variable d'intérêt.

4 Choisissez un échantillon pour la quantification réelle en incluant la plus petite unité dans le premier ansemble, puis la deuxième plus petite unité classée dans le deuxième ensemble et ainsi de suite, jusqu'à ce que la plus grande unité classée soit sélectionné dans le dernier ensemble.

5 Répétez les étapes 1 à 4 pour m cycles pour obtenir un échantillon de taille mk pour la quantification réelle. La taille, n = mk, est obtenue pour l'analyse.

Exemple

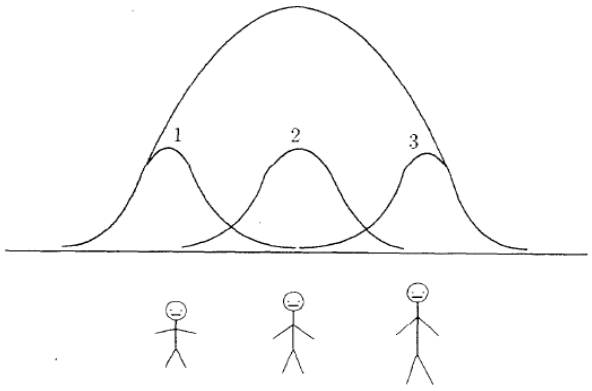

FIG. 1.1 – Distributions fréquentielles de hauteurs de différents rangs superposées sur la distribution de fréquence de la population de toutes les hauteurs - un diagramme schématique

1.3 Notation

Pour résoudre certaines idées considérons un échantillon aléatoire d'une distribution F, qui admet une fonction de densité $f(x)$,avec une moyenne μ et une variance σ^2. Par rapport à SRS, RSS utilise une unité, à savoir $X_{1(1:m)}$ l'unité la plus basse de cet ensemble ,puis $X_{2(2:m)}$ la deuxième unité la plus basse d'un autre ensemble indépendant de m unités, et enfin $X_{m(m:m)}$, la plus grande unité .Ce processus peut étre décrit à la figure 1.

Les dernière m unité $X_{1(1:m)}, X_{2(2:m),}...X_{m(m:m)}$ sont indépendants de manière indentique, et $X_{i(i:m)}$ est la statistique d'ordre i dans un échantillon aléatoire de taille m de distribution $F(x)$. Il est donc utile de comparer un RSS de taille m avec un SRS de mème taille m. De toute évidence, RSS serait un sérieux prétendant au SRS dans les cas où la tache d'assemblage des unités d'échantillonnage est facile et où leur classement relatif par rapport à la caractéristique étudiée peut étre effectué à un cout dérisoire.

TAB. 1.1 – Affichage de m^2 unité en m ensembles de taille m

$\left(X_{1(1:m)}\right)$	$X_{1(2:m)}$	...	$X_{1((m-1):m)}$	$X_{1(m:m)}$
$X_{2(1:m)}$	$\left(X_{2(2:m)}\right)$	...	$X_{2((m-1):m)}$	$X_{2(m:m)}$
.				
.				
.				
$X_{m(1:m)}$	$X_{m(2:m)}$	...	$X_{m((m-1):m)}$	$\left(X_{m(m:m)}\right)$

1.4 Echantillonnage statistiques

Dans cette section on va présenter quelques notions sur les différentes méthodes d'échantillonnage.

Définition 1.3 (Théorie de l'échantillonnage) *est un étude des liaisons existant entre une population et les échantillons de cette population, prélèvés par sondage.*

Définition 1.4 (Méthodes d'échantillonnage) *les méthodes d'échantillonnage sont des ensembles de méthodes permettant de réaliser un sondage (de prélever un échantillon de données) au sein d'une population, de manière à reproduire un échantillon aussi représentatif que possible de cette population.*

1.4.1 Evaluation de ces méthodes

le système d'échantillonnage sera jugé d'après la qualité des approximations des paramètres de la population, calculées sur l'échantillon prélevé . Pour cela, on étudiera la loi des caractéristiques classiques d'un échantillon (moyenne arithmétique , variance empirique,. . .).

1.4.2 Méthodes d'échantillonnages

1.4.3 Méthodes empiriques

les plus utilisées par les instituts de sondage. Leur précision ne peut pas être calculée et leur réussite dépend de l'expertise des enquêteurs.

Echantillonnage sur la base du jugement Echantillon prélevé à partir d'avis d'experts, qui connaissent bien la population et sont capable de dire quelles sont les entités représentatives.(par exemple, dans les campagnes électorales certains districts électoraux sont des indicateurs fiables de l'opinion publique). Pbme : l'avis des experts est subjectif.

Echantillonnage par la méthode des quotas : Echantillon prélevé librement à condition de respecter une composition donnée à l'avance (sexe, âge,. . .).

Pbme : repose sur la pertinence des catégories retenues.

1.4.4 Méthodes aléatoires

Reposent sur le tirage au hasard d'échantillons et sur le calcul des probabilités.

Echantillonnage aléatoire simple On prélève dans la population, des individus au hasard, sans remise : tous les individus ont la même probabilité d'être prélevés, et ils le sont indépendamment les uns des autres.

Echantillonnage aléatoire stratifié Supposont que la population soit stratifiée, i.e.constituée de sous-populations homogènes, les strates. (ex : stratfication

part ranche d'age). Dans chaque strate, on fait un échantillonnage aléatoire simple, de taille proportionnelle à la taille de strate dans la population (échantillon représentatif). Les individus de la population n'ont pas tous la

même probabilité d'être tirés. Nécessite une homogénéité des strates. Augmente la précision des estimations.

Echantillonnage par grappe On tire au hasard des grappes ou familles d'individus, et on examine tous les individus de la grappe (ex : on tire des immeubles puis on interroge tous les habitants). La méthode est d'autant meilleure que les grappes se ressemblent et que les individus d'une même grappe sont différents, contrairement aux strates.

1.5 Notations et formulation d'un échantillonnage classé

Soit F la fonction de répartition de la population d'intérèt, dont on suppose qu'elle admet une densité f. Soient $\mu \in R$ et $\sigma^2 > 0$ l'espérance et la variance de cette population. Admettons que l'on cherche à estimer μ par échantillonnage classé.

Par construction, un échantillon classé univarié (ETU) est formé de variables aléatoires $X_1^*, ..., X_k^*$ telles que X_i^* a la même distribution que la i^{eme} statistique d'ordre d'un échantillon aléatoire de taille k de loi F.

X_1^* représente la plus petite observation du premier échantillon, et donc que sa distribution est celle du minimum d'un échantillon de taille k, cette variable aléatoire ne constitue pas nécessairement le plus petit élément de l'échantillon classé.

Les variables aléatoires $X_1^*, ..., X_k^*$ composant un échantillon constituent collectivement ce que l'on appelle traditionnellement un cycle, en répétant le cycle à m reprises indépendantes, on obtient alors un échantillon classé de taille $k \times m$ observations. Les éléments de l'échantillon seront alors notés $X_{(1)j}^*, ..., X_{(k)j}^*$ où $j = 1, ..., m$.

Par exmple avec m est le nombre des cycle

$$
\begin{array}{rcl}
 & & Cycle\ 1 \\
X_{(1)11} & \leq & X_{(2)11} \leq X_{(3)11} \Longrightarrow X_{(1)1} \\
X_{(1)21} & \leq & X_{(2)21} \leq X_{(3)21} \Longrightarrow X_{(2)1} \\
X_{(1)31} & \leq & X_{(2)31} \leq X_{(3)31} \Longrightarrow X_{(3)1} \\
 & & Cycle2 \\
X_{(1)12} & \leq & X_{(2)12} \leq X_{(3)12} \Longrightarrow X_{(1)2} \\
X_{(1)22} & \leq & X_{(2)22} \leq X_{(3)22} \Longrightarrow X_{(2)2} \\
X_{(1)32} & \leq & X_{(2)32} \leq X_{(3)32} \Longrightarrow X_{(3)2} \\
 & & ..... \\
 & & ..... \\
 & & ..... \\
 & & Cycle\ m \\
X_{(1)1m} & \leq & X_{(2)1m} \leq X_{(3)1m} \Longrightarrow X_{(1)m} \\
X_{(1)2m} & \leq & X_{(2)2m} \leq X_{(3)2m} \Longrightarrow X_{(2)m} \\
X_{(1)3m} & \leq & X_{(2)3m} \leq X_{(3)3m} \Longrightarrow X_{(3)m}
\end{array}
$$

avec m est le nombre des cycle .

Pour la suite, notons par $X_1, ..., X_k$, les éléments d'un échantillon aléatoire simple (SRS) et par $X_{(1)} < ... < X_{(k)}$ les statistiques d'ordre qui y sont associées.

Les fonctions de densité d'un échantillon aléatoire simple SRS et d'un échantillon classé ordonné RSS sont respectivement données par

$$g_{SRS}(x_1, ..., x_k) = \prod_{i=1}^{k} f(x_i)$$

et

$$g_{RSS}(x_1, ..., x_k) = \prod_{i=1}^{k} f_i(x_i),$$

avec

$$f_i(x) = \frac{k!}{(i-1)!(k-i)!}\{F(x)\}^{i-1}\{1-F(x)\}^{k-i}f(x)$$

est la fonction de densité de la i^e statistique d'ordre d'un SRS de taille k de la loi F.

Lemme 1.1 *Soient $X_{(1)} < X_{(2)} < ... < X_{(k)}$ les statistiques d'ordre associées à un échantillon aléatoire de densité f. Soit f_i la densité de $X_{(i)}$. Alors* $\sum_{i=1}^{k} f_i(x) = kf(x), x \in \mathbb{R}$.

Preuve: Pour une valeur de x donnée et pour $p = F(x)$, on a par définition :

$$\begin{aligned}\sum_{i=1}^{k} f_i(x) &= kf(x) \sum_{i=1}^{k} \left(\begin{array}{c} k-1 \\ i-1 \end{array} \right) [F(x)]^{i-1}[1-F(x)]^{k-i} \\ &= kf(x) \sum_{j=0}^{k-1} \left(\begin{array}{c} k-1 \\ j \end{array} \right) p^j(1-p)^{k-1-j} \\ &= kf(x),\end{aligned}$$

puisque la somme représente la probabilité qu'une variable binomiale de paramètres $k-1$ et p prenne une valeur quelconque dans l'ensemble $\{0, 1, ..., k-1\}$. ■

1.6 Estimation des moyennes d'échantillonnale

Soit

$$\bar{X} = \frac{1}{mk} \sum_{j=1}^{m} \sum_{i=1}^{k} X_{ij}$$

la moyenne empirique de SRS (échantillon aléatoire simple). Il est bien connu que $\bar{X}$ est un estimateur sans biais de la moyenne μ de la population totale. Autrement dit,

$$E\left(\bar{X}\right) = \mu.$$

De mème pour un échantillon classé (RSS), on a

$$\bar{X}_{RSS} = \frac{1}{mk} \sum_{j=1}^{m} \sum_{i=1}^{k} X_{ij}$$

et donc on aura $E\left(\bar{X}_{RSS}\right) = \mu$.

1.6.1 Estimation de l'espérance d'une fonction transformée

Soit $h(X)$ une fonction quelconque de X et soit μ_h la moyenne de $h(X)$ c'est -à-dire $\mu_h = E(h(X))$.

Nous considérons dans cette séction l'estimation de μ_h en utilisont un échantillon d'ensemble classé.

Exemple 1.1
- *Soit $h(X) = x^l$, $l = 1, 2, ...$ correspondant à l'estimation des momonts de population .*
- *$h(X) = I\{x \leq c\}$ où $I\{i\}$ est la fonction indicatrice usuelle correspondant à l'estimation de la fonction de distribution .*
- *$h(X) = \frac{1}{\lambda} k\left(\frac{t-x}{\lambda}\right)$ où k est une fonction donnée et λ constante donnée correspondant à l'estimation de fonction de densité.*

Nous supposons que la variance de $h(X)$ existe.La moyenne est donnée comme suite :

$$\hat{\mu}_{RSS} = \frac{1}{km} \sum_{r=1}^{k} \sum_{i=1}^{m} X_{[r]i}$$

Nous considérons d'abord les propriétés statistiques de $\hat{\mu}_{RSS}$ et puis l'efficacité du RSS par rapport à SRS dans l'estimation des moyennes.

Tout d'abord, nous avons le résultat suivant.

Théorème 1.1 *Supposons que le mécanisme de classement dans RSS est consistent. Ensuite on a*

(i) *L'estimateur $\hat{\mu}_{RSS}$ est non biaisé, ie $E(\hat{\mu}_{RSS}) = \mu$.*

(ii) *$Var(\hat{\mu}_{RSS}) \leq \frac{\sigma^2}{mk}$ où σ^2 dénote la variance et l'inégalité est strict sauf si le mécanisme de classement est purement aléatoire.*

(iii) *Comme $m \to +\infty$*

$$\sqrt{mk}\,(\hat{\mu}_{RSS} - \mu) \to \mathcal{N}\left(0, \sigma^2_{RSS}\right),$$

dans la distribution, où,

$$\sigma^2_{RSS} = \frac{1}{k} \sum_{r=1}^{k} \sigma^2_{[r]}$$

ici $\sigma^2_{[r]}$ dénote la variance de $X_{[r]i}$.

Preuve:

(i) Il découle de l'égalité fondamentale que

$$\begin{aligned} E\left(\hat{\mu}_{RSS}\right) &= \frac{1}{mk}\sum_{r=1}^{k}\sum_{i=1}^{m}E\left(X_{[r]i}\right)=\frac{1}{k}\sum_{r=1}^{k}E\left(X_{[r]i}\right) \\ &= \frac{1}{k}\sum_{r=1}^{k}\int xdF_{[r]}\left(x\right)\xrightarrow{\mathcal{L}}\int xd\frac{1}{k}\sum_{r=1}^{k}F_{[r]}\left(x\right) \\ &= \int xdF\left(x\right)=\mu \end{aligned}$$

(ii)

$$\begin{aligned} var\left(\hat{\mu}_{RSS}\right) &= \frac{1}{\left(mk\right)^{2}}\sum_{r=1}^{k}\sum_{i=1}^{m}var\left(X_{[r]i}\right)=\frac{1}{mk^{2}}\sum_{r=1}^{k}var\left(X_{[r]}\right) \\ &= \frac{1}{mk}\left(\frac{1}{k}\sum_{r=1}^{k}\left(E\left(X_{[r]}\right)^{2}-\left[E\left(X_{[r]}\right)\right]^{2}\right)\right) \\ &= \frac{1}{mk}\left(m_{2}\ -\frac{1}{k}\sum_{r=1}^{k}\left[E\left(X_{r}\right)\right]^{2}\right), \end{aligned}$$

Où m_2 désigne le deuxième moment. Il découle de L'inégalité de Caushy-Schwarz

$$\frac{1}{k}\sum_{r=1}^{k}\left[E\left(X_{[r]}\right)\right]^{2}\geq\left(\frac{1}{k}\sum_{r=1}^{k}E\left(X_{[r]}\right)\right)^{2}=\mu^{2}$$

où l'égalité ne vaut que lorsque $E\left(X_{[1]}\right)=...=E\left(X_{[r]}\right)$ dans quel cas le mécanisme de classement est purement aléatoire.

(iii) Par l'égalité fondamentale $\mu=\frac{1}{k}\sum_{r=1}^{k}\mu_{[r]}$ où $\mu_{[r]}$ désigne la moyenne de $\left(X_{[r]i}\right)$. Ensuite, nous pouvons écrire

$$\begin{aligned} \sqrt{mk}\left(\hat{\mu}_{RSS}-\mu\right) &= \frac{1}{\sqrt{k}}\sum_{r=1}^{k}\sqrt{m}\left[\frac{1}{m}\sum_{i=1}^{m}\left(X_{[r]i}\right)-\mu_{[r]}\right] \\ &= \frac{1}{k}\sum_{r=1}^{k}Z_{mr} \end{aligned}$$

Par le théorème centrale limite multivariée, $(Z_{m1},...,Z_{mk})$ converge vers une distribution normale multivariée avec vecteur moyen nul et matrice de covariance donné par $Diag\left(\sigma_1^2,...,\sigma_k^2\right)$. Nous savons que σ^2/mk est la

variance de l'estimateur de moment en μ sur un échantillon aléatoire simple de taille mk. Le théorème implique que le moment de μ basé sur un échantillon RSS a toujours une variance inférieure à sa contrepartie basée sur un échantillon SRS de la même taille. Dans le contexte de RSS, nous avons tacitement supposé que le coût ou l'effort pour prélever des unités d'échantillonnage de la population et ensuite leur classement est négligeable. Quand on compare l'efficacité d'une procédure statistique basée sur un échantillon classé avec le procédure statistique basée sur un échantillon de aléatoire simple, nous supposons que les deux échantillons ont la même taille.

■

Théorème 1.2 *Soit $\hat{\mu}_{SRS}$ la moyenne d'un échantillon aléatoire simple de taille mk. on définit l'efficacité relative de RSS par rapport au SRS dans l'estimation de μ comme suit :*

$$RE\left(\hat{\mu}_{RSS}, \hat{\mu}_{SRS}\right) = \frac{var\left(\hat{\mu}_{SRS}\right)}{var\left(\hat{\mu}_{RSS}\right)}$$

Implique que $RE\left(\hat{\mu}_{RSS}, \hat{\mu}_{SRS}\right) \geq 1$.Pour étudier l'efficacité relative de manière plus détaillée, nous déduisons ce qui suit

$$\begin{aligned}
\sigma^2_{SRS} &= \frac{1}{k}\sum_{r=1}^{k}\sigma^2_{[r]} \\
&= \frac{1}{k}\sum_{r=1}^{k}\left(E\left(X_{[r]}\right)^2 - \left[E\left(X_{[r]}\right)\right]^2\right) \\
&= \frac{1}{k}\sum_{r=1}^{k}E\left(X_{[r]}\right)^2 - \mu^2 + \mu^2 - \frac{1}{k}\sum_{r=1}^{k}\left[E\left(X_{[r]}\right)\right]^2 \\
&= \sigma^2 - \frac{1}{k}\sum_{r=1}^{k}\left(\mu_{[r]} - \mu\right)^2
\end{aligned}$$

Ainsi,nous pouvons exprimer l'efficacité relative commes suit

$$RE\left(\hat{\mu}_{RSS}, \hat{\mu}_{SRS}\right) = \frac{\sigma^2}{\sigma^2_{RSS}} = \left[1 - \frac{\frac{1}{k}\sum_{r=1}^{k}\left(\mu_{[r]} - \mu\right)^2}{\sigma^2}\right]^{-1}.$$

Il ressort clairement de l'expression ci-dessus que tant qu'il existe au moins un r tel que $\mu_{[r]} \neq \mu$,l'efficacité relative est supérieure à 1.Pour une distribution sous-jacente donnée et une fonction h donnée, l'efficacité relative peut

être calculée au moins, en principe. Maintenent nous discutons de l'efficacité relative de manière plus détaillée pour le cas particulier où $h(x) = x$.

Sur la base des calculs effectués sur un certain nombre de distributions sous-jacentes, Mclntyre en 1996 a formulé la conjecture suivante :

l'efficacité relative de RSS par rapport au SRS dans l'estimation des variance est donnée comme suite :

TAB. 1.2 – L'efficacité relative de RSS avec $k = 2, 3, 4$ pour certaines distributions

Distribution	μ	σ^2	γ	k	2	3	4
$Uniforme$	0.500	0.083	0.000	1.80	1.500	2.000	2.500
$Exponentiele$	1.000	1.000	2.000	9.00	1.333	1.636	1.920
$Gamma(0.5)$	0.500	0.500	2.828	15.0	1.254	1.483	1.696
$Normale$	0.000	1.000	0.000	3.00	1.467	1.914	2.347

La moyenne de la population est compris entre 1 et $(k+1)/2$ où k est la taille de l'ensemble pour distributions sous-jacentes symétriques, l'efficacité relative n'est pas beaucoup moins que $(k+1)/2$ toutefois lorsque la distribution sous-jacente devient asymétrique,l'efficacité relative diminue.

Le tableau 1 ci-dessous est partiellement reproduit de tableau de Dell et Clutter (1972)[10]. Les notations μ, σ^2, γ et κ dans le tableau sont respectivement, la moyenne, la variance, l'asymétrie et le kurtosis.

1.7 Estimation de la variance d'échantillonnale classée

Nous allons maintenant intéressons à la variance de la moyenne échantillonnale classée.

Posons d'abord $\mu_i = E\left(X_{RSS(i)}\right)$ et $\sigma_i^2 = var\left(X_{RSS(i)}\right)$ pour tout $i \in \{1, ..., k\}$.

Puisque les variables $X^*_{(1)}, ..., X^*_{(k)}$ sont mutuellement indépendantes, nous

avons alors

$$\begin{aligned} var\left(\bar{X}_{RSS}\right) &= \frac{1}{k^2}\sum_{i=1}^{k}\sigma_i^2 \\ &= \frac{1}{k^2}\sum_{i=1}^{k}\int (x-\mu_i)^2 f_i(x)\,dx \\ &= \frac{1}{k^2}\sum_{i=1}^{k}\int (x-\mu+\mu-\mu_i)^2 f_i(x)\,dx \\ &= \frac{1}{k^2}\sum_{i=1}^{k}\int [(x-\mu)^2+2(\mu-\mu_i)(x-\mu)+(\mu-\mu_i)^2] f_i(x)\,dx \end{aligned}$$

en vertu de Lemma 1.1

$$\sum_{i=1}^{k}\int (x-\mu)^2 f_i(x)\,dx = \int (x-\mu)^2 kf(x)\,dx = kvar(X) = k\sigma^2$$

et

$$2\sum_{i=1}^{k}(\mu-\mu_i)\int (x-\mu) f_i(x)\,dx = -2\sum_{i=1}^{k}(\mu-\mu_i)^2$$

de sorte qu'au total,

$$\frac{1}{k^2}\sum_{i=1}^{k}\sigma_i^2 = \frac{1}{k^2}\left\{k\sigma^2\sum_{i=1}^{k}(\mu-\mu_i)^2\right\} = \frac{\sigma^2}{k} - \frac{1}{k^2}\sum_{i=1}^{k}(\mu-\mu_i)^2 \qquad (1.1)$$

et donc on a

$$var\left(\bar{X}_{RSS}\right) \leq \frac{\sigma^2}{k} = var\left(\bar{X}\right)$$

Cette inégalité revient à dire que l'estimateur de la moyenne fondé sur un échantillon classé est plus précis que celui qui est déduit d'un échantillon aléatoire simple.

Bien que la démonstration que nous venons de donner ne vaut que si le tri des observations se fait sans erreur, le résultat reste vrai à moins que les rangs servant à l'échantillonnage classé aient été attribués aléatoirement.

Dans le cas de tri aléatoire, on aura $\mu_1, \mu_2, .., \mu_k$ ce qui conduira alors à l'égalité des variances.

Takahasi & Wakimoto (1968)[24, Takahasi,] ont aussi démontré que

$$1 \leq eff\left(\bar{X}_{RSS}\backslash\bar{X}\right) \equiv \frac{var\left(\bar{X}\right)}{var\left(\bar{X}_{RSS}\right)} \leq \frac{1}{2}(k+1)$$

Où la borne supérieure est atteinte si et seulement si nous sommes en présence de la loi uniforme. Comme Dell & Clutter (1972)[10] l'ont fait remarquer, la valeur de $eff\left(\bar{X}_{RSS}\backslash\bar{X}\right)$ est d'ailleurs proche du maximum pour un grand nombre de lois unimodales.

– L'estimateur SRS de la variance de population σ^2 est donné par

$$s^2_{SRS} = \frac{1}{mk-1}\sum_{r=1}^{k}\sum_{i=1}^{m}\left(X_{ri-}\bar{X}_{SRS}\right)^2$$

– L'estimateur RSS de la variance de population σ^2 est donné par

$$s^2_{RSS} = \frac{1}{mk-1}\sum_{r=1}^{k}\sum_{i=1}^{m}\left(X_{[r]i}-\bar{X}_{RSS}\right)^2$$

– Les propriétés de s^2_{RSS} ont été étudiés par Stokes, et on a

$$E\left(s^2_{RSS}\right) = \sigma^2 + \frac{1}{k(mk-1)}\sum_{r=1}^{k}\left(\mu_{[r]}-\mu\right)^2$$

– Une mesure appropriée de l'efficacité relative de s^2_{RSS} par rapport à s^2_{SRS} est ensuite donnée par

$$\begin{aligned} RE\left(s^2_{RSS}, s^2_{SRS}\right) &= \frac{var\left(s^2_{SRS}\right)}{MSE\left(s^2_{RSS}\right)} \\ &= \frac{var\left(s^2_{SRS}\right)}{var\left(s^2_{RSS}\right) + \left[\frac{1}{k(km-1)}\sum_{r=1}^{k}\left(\mu_{[r]}-\mu\right)^2\right]^2} \end{aligned}$$

On peut facilement voir que

$$RE\left(s^2_{RSS}, s^2_{SRS}\right) < ARE\left(s^2_{RSS}, s^2_{SRS}\right).$$

Puisque

$$\frac{1}{k}\sum_{r=1}^{k}\left(\mu_{[r]}-\mu\right)^2 < \sigma^2,$$

Il est clair que $\frac{1}{k(mk-1)}\sum_{r=1}^{k}\left(\mu_{[r]}-\mu\right)^2$ diminuera si k ou m augmente .C'est-à-dire que les RE vont converger de plus en plus vers les ARE lorsque k ou m augmentent.

Fonction de densité $f(x) = (3/2)\,x^2 I\{-1 \le x \le 1\}$, et la distribution gamma a la fonction de densité $f(x) = x^4\exp(-x)/\Gamma(5)\,I\{x \ge 0\}$.

TAB. 1.3 – Efficacité relative $var\left(s^2_{SRS}\right)/MSE\left(s^2_{RSS}\right)$, pour $m = 1$ et $m \to \infty$

distribution	k	m=1	m$\to \infty$
(i) Gamma	2	0.71	1.02
	3	0.81	1.08
	4	0.91	1.16
	5	1.00	1.23
	6	1.09	1.35
(ii) uniforme	2	0.72	1.00
	3	0.92	1.11
	4	1.09	1.25
	5	1.20	1.40
(iii) Normal	2	0.68	1.00
	3	0.81	1.08
	4	0.93	1.18
	5	1.03	1.27

Il convient de noter que, dans l'estimation de la variance, RSS n'est pas nécessairement plus efficace que le SRS lorsque la taille de l'échantillon est petite et que l'efficacité est beaucoup plus faible que dans l'estimation de la moyenne de la population, même quand RSS est bénéfique.

RSS est le plus utile lorsque la moyenne et la variance de la population doivent être estimées.

C'est en effet une question naturelle de savoir si de meilleures estimations de σ^2 sur un échantillon RSS peut être trouvé ?.

Nous abordons cette question dans la sous-section suivante.

1.8 Variance minimale non biaisée estimation non négative

Nous démontrons dans cette sous-section qu'il est possible de construire une classe d'estimations impartiales non négatives de σ^2 basée sur un équilibre échantillon RSS, quelle que soit la nature de la distribution sous-jacente. À cette fin, nous avons besoin de l'identité de base suivante qui découle directement de

$$\sigma^2 = \frac{1}{k}\left[\sum_{r=1}^{k}\sigma^2_{[r]} + \sum_{r=1}^{k}\mu^2_{[r]}\right] - \mu^2$$

Rappeler que

$$\bar{X}_{i:RSS} = \frac{1}{k}\sum_{r=1}^{k} X_{[r]i}$$

fournit une estimation impartiale de la moyenne μ sur la base des données du $i^{éme}$ cycle d'un RSS.

$$W_i = \sum_{r=1}^{k} \left(X_{[r]i} - \bar{X}_{i:RSS}\right)^2, \qquad i = 1, ..., m.$$

De l'identité de base, il est clair qu'une estimation impartiale de σ^2 pouvent être obtenu en branchant des estimations impartiales de $\sigma^2_{[r]} + \mu^2_{[r]}$ et μ^2. Puisque $\sum_{i=1}^{m} X^2_{[r]i}/m$ est une estimation impartiale de l'ancien terme et $\bar{X}_{i:RSS}\bar{X}_{j:RSS}$ pour $i \neq j$ est une estimation impartiale de μ^2 il s'ensuit facilement qu'une estimation impartiale de σ^2 est donné par

$$\hat{\sigma}^2 = \frac{\sum_{r=1}^{k}\sum_{i=1}^{m}\left[X_{[r]i}\right]^2}{mk} - \frac{\sum_{i\neq j}\bar{X}_{i:RSS}\bar{X}_{j:RSS}}{m\left(m-1\right)}.$$

L'estimation ci-dessus peut être facilement simplifiée car

$$\hat{\sigma} = \frac{W}{mk} + \frac{B}{\left(m-1\right)k}$$

où W et B représentent, respectivement, l'entre-et à l'intérieur-somme de cycle de carrés de l'ensemble des données équilibrées, défini comme

$$B = k\sum_{i=1}^{m}\left(\bar{X}_{i:RSS} - \bar{X}_{RSS}\right)^2, W = \sum_{i=1}^{m} W_i$$

Il est évident que $\hat{\sigma}^2$ est non négatif

$$\begin{aligned} X_{[r]} &= \left(X_{[r]1}, ..., X_{[r]m}\right)^t, r = 1, ..., k, \\ X &= \left(X^t_{[1]}, ..., X^t_{[k]}\right)^t \end{aligned}$$

1.9 Taille d'échantillon classé

Toute personne voulant faire une étude statistique concerne le choix de la taille de l'échantillon.

D'une part, comme chaque donnée mesurée amène une information supplémentaire du fait de son rang parmi les k unités de son échantillon, il est

évident que plus k est élevé, plus nous obtiendrons d'information additionnelle si tous les rangs sont attribués de façon exacte.

D'autre part, plus la taille est grande, plus il sera difficile d'établir un ordre qui, faut-il le rappeler, est déterminé sans prendre une seule mesure sur les observations. Ainsi, le risque d'erreur augmente à mesure que croit la taille de l'échantillon classée. En somme, il faut ètre capable de trouver un certain équilibre entre ces deux aspects du problème.

Par ailleurs, les contraintes monétaires entrent aussi en ligne de compte lorsqu'il faut déterminer la taille d'un échantillon ou d'un échantillon classé. Afin de pouvoir fixer la taille optimale, il faudra donc ètre capable d'estimer les probabilités de commettre des erreurs dans les rangs ainsi que d'avoir une bonne idée de l'impact de ces erreurs potentielles sur les procédures statistiques qui seront utilisées ultérieurement.

Enfin, notons qu'il est parfois avantageux de recourir à une forme d'échantillonnage classé non équilibré, c'est-à-dire dans laquelle chacune des statistiques d'ordre n'apparait pas obligatoirement une seule fois dans l'échantillon classé $X_1^*, ..., X_k^*$.En effet, prenons le cas où nous avons une distribution unimodale et symétrique autour de la médiane.

Supposons que nous désirons faire de l'inférence sur cette médiane à l'aide d'un échantillon classé de taille k impaire. Dans cette situation, il serait adéquat de prendre la médiane $X_{\frac{k+1}{2}}$ de chacun des k échantillons afin de batir l'échantillon classé en question. Toutefois, cette problématique ne sera pas considérée plus avant dans cet essai.

1.10 Estimation de la fonction de répartition

Stokes & Sager 1988[17] ont montré comment estimer une fonction de répartition à partir de l'information supplémentaire qu'apporte l'échantillonnage classé. Supposons que nous disposons de $X_{(1)j}^*, ..., X_{(k)j}^*$ pour tout $j \in \{1, ..., m\}$. En d'autres termes, supposons que nous avons en main un échantillon classé de taille k et de m cycles tiré d'une population de loi F.

L'estimation proposée par Stokes & Sager en 1988 est donnée par

$$F^*(t) = \frac{1}{mk} \sum_{i=1}^{k} \sum_{j=1}^{m} I_{(-\infty,t)} \left(X_{(i)j}^* \right). \tag{1.2}$$

Ces auteurs ont montré que F^*est une estimation non biaisée de F.De plus, ils ont établi que si $\hat{F}$ est la fonction de répartition empirique d'un échantillon aléatoire simple de taille mk, alors

$$var\left\{F^*(t)\right\} \leq var\left\{\hat{F}(t)\right\} \ pour\ tout\ t \in \mathbb{R}$$

En d'autres termes, l'estimation de F extraite d'un échantillon classé est à la fois sans biais et plus précise, en moyenne, que celle déduite d'un échantillon aléatoire simple par la méthode classique.

Chapitre 2

Quantiles par RSS

2.1 Introduction

Les quantiles ont été introduits en statistique descriptive pour désigner des valeurs d'un caractère quantitatif repartissant la population étudiée en classes d'effectifs égaux. on dit médiane plutôt que quantile de niveau 0 ; 5, premier quartile plutôt que quantile de niveau 0 ; 25 et troisième quartile plutôt que quantile de niveau 0 ; 75. Les quantiles dont les niveaux sont des multiples de un dixième (resp. un centième) sont appelés déciles (resp. centiles).

En probabilité, la notion de quantile s'applique de manière tout-à-fait générale aux variables aléatoires à valeurs réelles, et elle permet de définir une fonction de répartition inverse (ou fonction quantile).

Comme on le verra plus loin, le procédé le plus courant de simulation d'une variable aléatoire de loi donnée s'appuie cette fonction.

2.2 Généralités

Pour une variable X, la fonction quantile se définit à partir de l'inverse de sa fonction de répartition. Quand cette fonction de répartition est strictement croissante, son inverse est défini sans ambiguïté. Mais une fonction de répartition constante sur tout intervalle dans lequel la variable aléatoire ne peut pas prendre de valeurs. De manière générale, soit F la fonction de répartition de la variable X.

Définition 2.1 *On appelle fonction quantile d'ordre p de x la fonction qui, à $p \in]0,1[$, associe*

$$x_p = F^{-1}(p) = \inf\{x : F(x) \geqslant p\},$$

où F^{-1} est souvent appelée l'inverse généralisé de F.

Remarque 2.1 *Pour certaines valeurs de p, on donne un nom particulier aux quantiles ; par exemple, pour $p = 0.5$ le quantile appelé médiane, pour $p = 0.25,\ 0.75$ le quantile appelé quartile, pour $p = 0.1, ..., 0.9$ le quantile appelé décile et pour $p = 0.01, ..., 0.09$ le quantile appelé centile,..., etc.*

2.3 Estimation de quantile

Cette section est consacrée à l'estimation des quantiles d'échantillon d'ensemble classé. Nous donnons d'abord la définition d'ensembles classés, analogues aux simples quantiles d'échantillons aléatoires, et examinons leurs propriétés. Nous examinons ensuit les prcédures d'inférence telles que la construction d'intervalles de confiance et vérification d'hypothéses pour les quantiles. Enfin, nous comparons les estimations quantiles RSS et les estimations quantiles SRS en termes de variance asymptotique.

2.3.1 Notations

La fonction de distribution cumulée (FDC) et la fonction de densité de probabilité (FDP) de la distribution sous-jacente sont désignées par F et f. Bien que, dans la plupart des cas, nous supposons que le classement des jugements dans RSS était parfait, nous permettons également la possibilité d'un classement imparfait dans certains des résultats discutés dans cet article. nous désignons par $X_{(r)}$ la statistique d'ordre classé lorsque le classement est parfait (pas d'erreur de classement ou de problèmes). et par $X_{[r]}$ lorsqu'il est possible que le classement soit imparfait. Les FDP et FDC de la statistique de la commande d'ordre classée du jugement dans un ensemble de taille k sont désignés par $F_{(x)}$ et $f_{(x)}$ si le classement est parfait, et par $F_{[x]}$ et $f_{[x]}$ si le classement peut être imparfait.

2.3.2 Propriétés des distributions d'un échantillons classé

On rappelle que la fonction de distribution empirique des ensembles classés est définie comme suite :

$$\hat{F}_{RSS}(x) = \frac{1}{mk} \sum_{r=1}^{k} \sum_{i=1}^{k} I\left\{X_{[r]} \leq x\right\}.$$

soit $n = mk$ et $I\{A\}$ est la fonction indicatrice sur l'ensemble A.

Proposition 2.1 *Pour $0 < p < 1$, le quantile empirique d'un échantillon classé, noté par $\hat{x}_n(p)$, est ensuite défini par,*

$$\hat{x}_n(p) = \inf\left\{x : \hat{F}_{RSS}(x) \geqslant p\right\}. \tag{2.1}$$

Théorème 2.1 *Supposons que le mécanisme de classement dans RSS soit cohérent. Puis,avec la probabilité 1,*

$$|\hat{x}_n(p) - x(p)| = \frac{2(\log n)^2}{f(x(p))\, n^{1/2}}, \tag{2.2}$$

pour tout n suffisamment grand.

Théorème 2.2 *Supposons que le mécanisme de classement dans RSS soit cohérent et que la fonction de densité f soit continue en $x(p)$ et positive dans un voisinage de $x(p)$. Alors,*

$$\hat{x}_n(p) = x(p) + \frac{p - \hat{F}_{RSS}(x(p))}{f(x(p))} + R_{n,} \tag{2.3}$$

avec probabilité 1, où

$$R_n = o\left(n^{-3/4}(\log n)^{3/4}\right),$$

lorsque $n \to \infty$.

La représentation de Bahadur suit immédiatement la normalité asymptotique du quantile d'échantillon classé.

Théorème 2.3 *Supposons que nous avons les mêmes conditions que dans le théorème 2.2 soient remplies.*

$$\sqrt{n}\left(\hat{x}_n(p) - x(p)\right) \to \mathcal{N}\left(0, \frac{\sigma^2_{k,p}}{f^2(x(p))}\right),$$

où,

$$\sigma^2_{k,p} = \frac{1}{k}\sum_{r=1}^{k} F_{[r]}(x(p))\left[1 - F_{[r]}(x(p))\right].$$

En particulier, si le classement est parfait, notant que $F_{[r]}(x(p)) = \mathcal{B}(r, k+r-1)$ dans ce cas,on a la formule suivante

$$\sigma^2_{k,p} = \frac{1}{k}\sum_{r=1}^{k} \mathcal{B}(r, k+r-1)\left[1 - \mathcal{B}(r, k+r-1)\right],$$

où $\mathcal{B}(r,s)$ désigne la fonction de distribution de bêta avec les paramètres r et s.

Théorème 2.4 *Soit* $0 < p_1 < ... < p_j < ... < p_l < 1$ *des probabilités. Soit* $\zeta = (x(p_1), ..., x(p_l))^T$ *et* $\hat{\zeta} = (\hat{x}(p_1), ..., \hat{x}(p_l))_T$.

Où $\sqrt{n}\left(\zeta - \hat{\zeta}\right) \to \mathcal{N}_l(0, \Sigma)$ dans la distribution

$$\Sigma = (\sigma_{ij}) = \frac{\frac{1}{k}\sum_{r=1}^{k} F_{[r]}(x(p_i))\left(1 - F_{[r]}(x(p_j))\right)}{[f(x(p_i))\, f(x(p_j))]}.$$

Définition 2.2 *Nous dérivons également certaines propriétés des statistiques sur les ordres classés. Une statistique d'ordre* $Z_{(k_n:n)}$ *est dite centrale si* $\frac{k_n}{n}$ *converge vers certains* p *tels que* $0 < P < 1$ *lorsque* $n \to \infty$.

Théorème 2.5 *Pour les statistiques sur les ordres classés par ordre central, nous avons l'analogue suivant des résultats pour les statistiques sur les ordres d'échantillonnage aléatoires simples :*

Théorème 2.6 **(i)** *si* $\frac{k_n}{n} = p + o\left(n^{-1/2}\right)$ *alors*

$$Z_{(k_n:n)} = x(p) + \frac{\frac{k_n}{n} - \hat{F}_{RSS}(x(p))}{f(x(p))} + R_n,$$

Avec probabilité 1,

$$R_n = o\left(n^{-3/4}(\log n)^{3/4}\right) \quad n \to \infty,$$

(ii) *Si*

$$\frac{k_n}{n} = p + \frac{c}{n^{1/2}} + o\left(n^{-1/2}\right)$$

Alors

$$\sqrt{n}\left(\left|Z_{(k_n:n)} - \hat{x}(p)\right|\right) \to \frac{c}{f(x(p))}$$

Avec un probabilité 1, *et* $\sqrt{n}\left(\left|Z_{(k_n:n)} - \hat{x}(p)\right|\right) \to \mathcal{N}\left(\frac{c}{f(x(p))}, \frac{\sigma^2_{k,p}}{f^2(x(p))}\right)$.

au sens de distribution.

2.4 Procédures d'inférence pour les quantiles de population basées sur un échantillon d'ensemble classé

Définition 2.3 *La méthode de noyau est l'une des méthodes d'estimation non paramétrique la plus utilisée. Rosenblatt (1956)[27], suivi de Parzen (1962)[15], ont proposé une classe d'estimateurs à noyau d'une densité de probabilité. Cet estimateur est une fonction de deux paramétres : Le noyau k et le paramétre de lissage h. Le succsè rencontré par cet estimateur èxplique par sa simplicité, sa flexibilité et aussi ses propriétés de convergence. Il laisse à l'utilisateur une grande latitude non seulement dans le choix du noyau k, mais aussi dansle choix du paramétre de lissage h.*

On appelle estimateur à noyau k de f, l'estimateur donné par :

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} k\left(\frac{x - x_i}{h}\right),$$

Où n est la taille de l'échantillon X, k est la fonction noyau et h est dit paramètre de lissage qui détermine son étendue.

L'estimateur à noyau de $F(x)$ est donné par :

$$F(x) = \int_{-\infty}^{x} \hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} k\left(\frac{x - x_i}{h}\right),$$

où $\int_{-\infty}^{+\infty} k(t)\,dt = 1$,
$\int_{-\infty}^{+\infty} k(t)\,dt = K(x)$.

Les résultats de la section 2.3.2 sont appliqués dans cette section pour les procédures d'inférence sur des quantiles tels que les intervalles de confiance et les tests d'hypothèses.

(i) Intervalle de confiance basé sur les statistiques des ordres classés. Pour construire un intervalle de confiance de coefficient $1-2\alpha$ pour $x(p)$, on cherche deux entiers ℓ_1 et ℓ_2 tels que $0 \leqslant \ell_1 < \ell_2 \leq n$ et que

$$P(Z_{(\ell_1:n)} \leq x(p)) \leq Z_{(\ell_2:n)} = 1 - 2\alpha.$$

Nous limitons notre attention aux intervalles de probabilité égale, c'est-à-dire intervalles satisfaisan

$$P\left(Z_{(\ell_1:n)} \leq x(p)\right) = 1 - \alpha,\ P\left(Z_{(\ell_2:n)} \leq x(p)\right) = \alpha$$

Ensuite, les entiers $\ell_1$et ℓ_2 peuvent être trouvés comme suit.

Soit N_r le nombre de $X_{[r]}$ avec r fixés qui sont inférieurs ou égaux à $X(p)$. Soit $N = \sum_{r=1}^{k} N_r$. Nous avons

$$P(Z_{(\ell_1:n)} \leq x(p)) = P(N \geqslant \ell_1).$$

Théorème 2.7 *Notons que N_r sont des variables aléatoires binomiales indépendantes avec $N_r \sim \mathcal{B}_i(m, p_r)$ où $p_r = F_{[r]}(x(p)) = P(N \geqslant \ell_1)$. Donc*

$$P(N \geqslant \ell_1) = \sum_{j=\ell_1}^{n} \sum_{(j)}^{k} \prod_{r=1}^{k} \binom{m}{i_r} p_r^{i_r} (1-p_r)^{m-i_r},$$

Où la somme Σ_j est sur tous les k d'entiers $(i_1, ..., i_k)$ satisfaisant $\sum_{r=1}^{n} i_r = j$. Alors, on peut déterminer ℓ_1 de telle sorte que la somme à la droite de l'égalité ci-dessus soit égale à ou proche de $1-\alpha$. De même ℓ_2 peut être déteminé. Bien que pas impossible, le calcul sera extrêmement lourd.Cependant, Lorsque n est grand, ℓ_1 et ℓ_2 peuvent être déterminés approximativement comme démontré ci-dessous. Note que

$$\begin{aligned} E(N) &= \sum_{r=1}^{k} mF_{[r]}(x(p)) = mkF(x(p)) = np, \\ Var(N) &= \sum_{r=1}^{k} mF_{[r]}(x(p)) \left(1 - mF_{[r]}(x(p))\right). \end{aligned}$$

D'après le théorème centrale limite , nous avons approximativement ,

$$\frac{N - np}{\sqrt{\sum_{r=1}^{k} mF_{[r]}(x(p)) \left(1 - mF_{[r]}(x(p))\right)}} \sim \mathcal{N}(0,1).$$

Où

$$\begin{aligned} \ell_1 &\simeq np - z_\alpha \sqrt{\sum_{r=1}^{k} mF_{[r]}(x(p)) \left(1 - mF_{[r]}(x(p))\right)}, \\ \ell_2 &\simeq np + z_\alpha \sqrt{\sum_{r=1}^{k} mF_{[r]}(x(p)) \left(1 - mF_{[r]}(x(p))\right)}, \end{aligned}$$

Où z_α désigne le $(1-\alpha)$ ème quantile de la distribution normale standard. Lorsque le classement est parfait, $F_{[r]}(x(p)) = B(r, k+r-1, p)$, et les

intervalles ci dessus peuvent être complètement déterminés. Cependant, en général $F_{[r]}(x(p))$ *est inconnu et doit être estimé. On peut prendre l'estimation comme étant* $F_{[r]}(x) = \frac{1}{m}\sum_{i=1}^{m} I\left\{X_{[r]i} \leq x\right\}$ *.Pour référence ultérieure, l'intervalle* $\left[Z_{(l_1:n)}, Z_{(l_2:n)}\right]$ *est désigné par* $\bar{I}_{s_n}$.

(ii) *Intervalle de confiance basé sur des quantiles d'échantillon classés. En utilisant le théorème 2.3, il est possible de construire un autre intervalle de confiance asymptotique du coefficient* $1-2\alpha$ *pour* $x(p)$ *peut être construit comme suit :*

$$\left[\hat{x}_n(p) - \frac{z_\alpha}{\sqrt{n}}\frac{\sigma_{k,p}}{f(x(p))}, \hat{x}_n(p) + \frac{z_\alpha}{\sqrt{n}}\frac{\sigma_{k,p}}{f(x(p))}\right].$$

Cet intervalle est noté $\bar{I}_{Q_n}$. *Puisque* $\bar{I}_{Q_n}$ *implique la quantité inconnue* $f(x(p))$, *nous devons la remplacer par un estimation cohérente dans la pratique. Dans la section suivante, nous examinerons l'estimation de* f *par la méthode du noyau en utilisant des données RSS. L'estimation de* f *par le noyau RSS peut très bien servir l'objectif ici.*

soit $\hat{f}_{RSS}$ *,l'estimation du noyau RSS de* f. *Alors, dans* $\bar{I}_{Q_n}$*,l'inconnu* $f(x(p))$ *peut être remplacé par* $\hat{f}_{RSS}(\hat{x}_n(p))$. *Note que les intervalles* $\bar{I}_{Q_n}$ *et* $\bar{I}_{s_n}$ *sont équivalents en ce sens que les deux intervalles se chevauchant approximativement alors que les coefficients de confiance sont les mêmes. Il résulte des théorèmes 2 et 5 que,*

avec la probabilité 1,

$$Z_{(\ell_1:n)} - \left[\hat{x}_n(p) - \frac{z_\alpha}{\sqrt{n}}\frac{\sigma_{k,p}}{f(x(p))}\right] = o\left(n^{-1/2}\right),$$

et

$$Z_{(\ell_2:n)} - \left[\hat{x}_n(p) - \frac{z_\alpha}{\sqrt{n}}\frac{\sigma_{k,p}}{f(x(p))}\right] = o\left(n^{-1/2}\right).$$

Notant que la longueur des deux intervalles est d'ordre $o\left(n^{-1/2}\right)$, *l'équivalence est établie. En pratique, l'un ou l'autre de ces deux intervalles pourrait être utilisé.*

(iii) *Tester d'hypothèses utilisant des quantiles d'échantillons classés. La normalité commune asymptotique des quantiles d'échantillon classés, comme indiqué dans le théorème 2.4, peut être utilisée pour tester des hypothèses impliquant des quantiles de population. Supposons que l'hypothèse nulle soit de la forme* $l^T\zeta = c$, *où* $\zeta = (x(p_1), ..., x(p_q))^T$ *est un vecteur de quantiles d'échantillons classés, et* l *et* c *sont donnés comme*

vecteur et scalaire de constantes, respectivement. La statistique de test peut alors être formée comme suit :

$$S_n = \frac{\sqrt{n}\left[l^T\hat{\zeta} - c\right]}{\sqrt{l^T\hat{\Sigma}l}},$$

Où $\hat{\Sigma}$ *est la matrice de covariance estimée donnée par*

$$\hat{\sigma}_{ij} = \frac{1}{k}\sum_{r=1}^{k} p_{ir}\left(1 - p_{ir}\right) / \left[\hat{f}\left(\hat{x}_n\left(p_i\right)\right)\hat{f}\left(\hat{x}_n\left(p_j\right)\right)\right].$$

Selon le théorème 2.4, la statistique de test suit asymptotiquement la distribution normale standard sous l'hypothèse nulle. la règle de décision peut être prise en conséquence.

2.5 Efficacité relative de l'estimation du quantile RSS par rapport à l'estimation du quantile SRS

Dans cette section, nous discutons de l'efficacité relative asymptotique (ARE) de l'estimation du quantile RSS par rapport au l'estimation du quantile SRS. L'estimateur de $\hat{x}_n\left(p\right)$ est le quantile d'échantillon $\hat{\zeta}_{np}$ d'un échantillon aléatoire simple de taille n. Il peut être trouvé à partir de n'importe quel manuel standard que $\hat{\zeta}_{np}$ a une distribution normale asymptotique avec la moyenne $x(p)$ et la variance $\frac{p(1-P)}{nf(x(p))}$. (Voir, par exemple, Serfling (1980[18], chapitre 2)).

Par conséquent, les ARE de $\hat{x}_n\left(p\right)$ en ce qui concerne à $\hat{\zeta}_{np}$ est donné par

voir Chen (2000)[7]. Par conséquent, l'efficacités relative asymptotique (ARE) donné par

$$ARE\left(\hat{x}_n\left(p\right), \hat{\zeta}_{np}\right) = \frac{p\left(1-P\right)}{\frac{1}{K}\sum_{r=1}^{k} F_{[r]}\left(x\left(p\right)\right)\left(1 - F_{[r]}\left(x\left(p\right)\right)\right)},$$

avec

$$ARE\left(\hat{x}_n\left(p\right), \hat{\zeta}_{np}\right) > 1$$

$ARE\left(\hat{x}_n\left(p\right), \hat{\zeta}_{np}\right)$ est toujours supérieur à 1 pour tout p, la quantité peut être très différente pour différentes valeurs de p. Pour mieux comprendre la nature de l'ARE, considérons le cas du classement parfait. Dans ce cas,

$$F_{[r]}\left(x\left(p\right)\right) = \mathcal{B}\left(r, k + r - 1, p\right).$$

2.5.1 Estimation de la fonction de densité avec un échantillon d'ensemble classé

Dans le contexte de RSS le besoin d'estimation de densité apparaît dans certaines procédures statistiques. Par exemple, l'intervalle de confiance et le test d'hypothèse. Les procédures basées sur des quantiles d'échantillons classés examinés dans la section 2.3 doivent disposer d'une estimation des valeurs de la fonction de densité pour certains quantiles. D'autre part, l'estimation de la densité a son propre intérêt indépendant. Une estimation de la densité peut révéler des caractéristiques importantes telles que l'asymétrie et la multimodalité de la distribution sous-jacente. Une estimation de la densité est un outil idéal pour la présentation des données aux clients afin de fournir des explications et des illustrations des conclusions qui ont été obtenues. Dans cette section, nous abordons la tâche de développer des méthodes d'estimation de la densité à l'aide de données RSS. La section 2.5.2 donne l'estimation de la fonction de densité f et examine ses propriétés. La section 2.5.3 traite de l'eficacité relative de l'estimation de la densité à l'aide des données RSS par rapport à son homologue dans le SRS.

2.5.2 L'estimation de la densité RSS et ses propriétés

Il existe une vaste littérature sur l'estimation de la densité dans le SRS. Une variété de méthodes ont été proposées et dévelopées, notamment le plus proche voisin, le noyau, la vraisemblance pénalisée maximale et la méthode du noyau adaptatif, etc.. Une bonne référence sur la méthodologie générale d'estimation de la densité est Silverman (1986)[19] . Chacune des différentes méthodes a ses propres avantages et inconvénients. Il n'existe pas d'accord universel sur la méthode à utiliser. Nous concentrerons notre attention sur la méthode de noyau pour l'estimation de la densité et ses ramifications. Nous avons choisi de traiter la méthode du noyau en partie parce que c'est un bon choix problèmes d'échantillons d'ensembles classés et en partie parce que ses propriétés mathématiques sont bien comprises dans de nombreuses applications pratiques. Pour faciliter la discussion qui suit, avant de traiter des propriétés de l'estimation RSS, nous donnons ci-dessous la définition de l'estimation SRS, ainsi que certaines de ses propriétés. Sur la base d'un échantillon aléatoire simple $X_1, ..., X_n$, l'estimation du noyau de f est donnée par

$$\hat{f}_{SRS}(x) = \frac{1}{nh} \sum_{i=1}^{n} k\left(\frac{x - X_i}{h}\right).$$

la moyenne et la variance de $\hat{f}_{SRS}(x)$ peuvent être

facilement dérivées

$$E(\hat{f}_{SRS}(x)) = \int \frac{1}{h}\sum_{i=1}^{n} k\left(\frac{x-X_i}{h}\right) f(t)\,dt, \tag{2.4}$$

$$Var\left(\hat{f}_{SRS}(x)\right) = \frac{1}{n}\int \frac{1}{h^2}\sum_{i=1}^{n} k\left(\frac{x-t}{h}\right)^2 f(t)\,dt \tag{2.5}$$

$$-\frac{1}{n}\left[\int \frac{1}{h}\sum_{i=1}^{n} k\left(\frac{x-t}{h}\right) f(t)\,dt\right]^2 \tag{2.6}$$

Pour motiver la définition de l'estimation avec les données RSS, notons que, du point de vue de l'égalité fondamentale, nous avons

$$f(x) = \frac{1}{k}\sum_{r=1}^{n} f(x)_{[r]}, \tag{2.7}$$

où $f_{[r]}$ désigne la fonction de densité correspondant à $F_{[r]}$. Le sous échantillon $X_{[r]i}$, $i = 1, ..., m$, est en effet un échantillon aléatoire simple issu de la distribution, $f_{[r]}$ on peut donc estimer la méthode du noyau en utilisant le sous-échantillon.

L'estimation du noyau $\hat{f}_{[r]}$ de $f_{[r]}$ en x sur la base du sous-échantillon est définie comme

$$\hat{f}_{[r]}(x) = \frac{1}{mh}\sum_{i=1}^{m} k\left(\frac{x-X_{[r]i}}{h}\right), \tag{2.8}$$

où K est une fonction du noyau et h est la bande passante à déterminer. Ainsi, une définition naturelle de l'estimation du noyau de f est donnée par

$$\hat{f}_{RSS}(x) = \frac{1}{k}\sum_{r=1}^{n} \hat{f}_{[r]}(x) = \frac{1}{kmh}\sum_{r=1}^{k}\sum_{i=1}^{m} k\left(\frac{x-X_{[r]i}}{h}\right) \tag{2.9}$$

si découle de (2.7) cela

$$\begin{aligned} E\left(\hat{f}_{RSS}(x)\right) &= \frac{1}{kh}\sum_{r=1}^{k} E\left(k\left(\frac{x-X_{[r]i}}{h}\right)\right) \\ &= \frac{1}{k}\sum_{r=1}^{k}\int k\left(\frac{x-t}{h}\right) f_{[r]}(t)\,dt \\ &= \int \frac{1}{h}\int k\left(\frac{x-t}{h}\right) f(t)\,dt = E\left(\hat{f}_{SRS}(x)\right), \end{aligned} \tag{2.10}$$

$$
\begin{aligned}
Var\left(f_{RSS}\left(x\right)\right) &= \frac{1}{mk^2}\sum_{r=1}^{K} Var\frac{1}{h}K\left(\frac{x-X_{[r]}}{h}\right) \\
&= \frac{1}{mk^2}\sum_{r=1}^{k}\left\{E\left[\frac{1}{h}K\left(\frac{x-X_{|r|}}{h}\right)^2\right]-\left[E\frac{1}{h}K\left(\frac{x-X_{|r|}}{h}\right)\right]^2\right\} \\
&= \frac{1}{mk}\left\{E\left[\frac{1}{h}K\left(\frac{x-X}{h}\right)\right]^2-\frac{1}{K}\sum_{r=1}^{K}\left[E\frac{1}{h}K\left(\frac{x-X_{|r|}}{h}\right)\right]^2\right\} \\
&= Var\left(\hat{f}_{SRS}\left(x\right)\right) \\
&\quad +\frac{1}{mk}\left\{\left[E\frac{1}{h}K\left(\frac{x-X\,|r|}{h}\right)\right]^2-\frac{1}{K}\sum_{r=1}^{K}\left[E\frac{1}{h}K\left(\frac{x-X_{|r|}}{h}\right)^2\right]\right\}
\end{aligned}
$$

si découle encore de (2.7) que

$$
E\left[\frac{1}{h}k\left(\frac{x-X}{h}\right)\right]=\frac{1}{k}\sum_{r=1}^{k}E\left[\frac{1}{h}k\left(\frac{x-X_{[r]}}{h}\right)\right].
$$

Par l'inégalité de Cauchy-Schwarz, nous avons

$$
E\left[\frac{1}{h}k\left(\frac{x-X}{h}\right)\right]^2<\frac{1}{k}\sum_{r=1}^{k}E\left[\frac{1}{h}k\left(\frac{x-X_{[r]}}{h}\right)\right]^2.
$$

En résumant l'argument ci-dessus, nous concluons que $\hat{f}_{RSS}\left(x\right)$ à les mêmes attentes que $\hat{f}_{SRS}\left(x\right)$ et une variance inférieure à celle de $\hat{f}_{SRS}\left(x\right)$. Cela implique que l'estimation RSS a une erreur quadratique moyenne intégrée ($MISE$) inférieure à l'estimation SRS. Le MISE d'une estimation $\hat{f}$ de f est défini par

$$
MISE\left(\hat{f}\right)=E\left(\int\left[\hat{f}\left(x\right)-f\left(x\right)\right]^2dx\right).
$$

La conclusion est valable si le classement est parfait ou non. Dans ce qui suit, nous supposons que f à certaines dérivées et que K satisfait les conditions :

i) k est symétrique

ii) $\int k\left(t\right)dt=1$ et $\int t^2k\left(t\right)dt\neq 0$.

Lemme 2.1 *Sous les hypothéses ci-dessus sur f et k tiré, comme $h \to 0$,*

$$\left[E\left(\frac{1}{h}k\left(\frac{x-X}{h}\right)\right)\right]^2 - \frac{1}{k}\sum_{r=1}^{k} E\left[\frac{1}{h}k\left(\frac{x-X_{[r]}}{h}\right)\right]^2 = \left[f^2(x) - \frac{1}{k}\sum_{r=1}^{k} f_{[r]}^2\right] + O\left(h^2\right).$$

Le lemme 2.1 peut être prouvé par un calcul simple permettant des développements de Taylor de f et $f_{[r]}$.

on oura

$$\Delta(f,k) = \int \left[\frac{1}{k}\sum_{r=1}^{k} f_{[r]}^2(x) - f^2(x)\right] dx.$$

Note que $\Delta(f,k) > 0$. Nous avons les résultats suivants.

Théorème 2.8 *Supposons que la même bande passante soit utilisée à la fois dans $\hat{f}_{SRS}(x)$ et $\hat{f}_{RSS}(x)$. Ensuite, pour k tiré et grand n,*

$$MISE\left(\hat{f}_{RSS}\right) = MSIE\left(\hat{f}_{SRS}\right) - \frac{1}{n}\Delta(f,k) + O\left(\frac{h^2}{n}\right).$$

nous examinons maintenant le cas particulier du classement parfait et obtenons des résultats asymptotiques qui éclairent les propriété de l'estimation de la densité RSS.lorsque classement est parfait $f_{[r]} = f(x)$, la statistique d'ordre 1.Preliérement, nous avons

Lemme 2.2 *Si, pour $r = 1, ..., k$, $f(r)$ est la fonction de densité de la statistique d'ordre r d'un échontillon de taille k à partir d'un distribution de fonction de densité f;nous avons la représentation suivante*

$$\frac{1}{k}\sum_{r=1}^{k} f_{(r)}^2(x) = k f_{(r)}^2 P(Y = Z),$$

Lemme 2.3 *oùY et Z sont indépendants avec la même distribution binomiale $B(k-1, F(x))$.*

En outre,

$$P(Y = Z) = \frac{1}{\sqrt{4\pi k F(x)[1 - F(x)]}} + o\left(\frac{1}{k}\right),$$

Preuve: *lorsque le classement est parfait, nous avons*

$$f_{(r)}(x) = \frac{k!}{(r-1)!(k-r)!} F^{r-1}(x)[1 - F(x)]^{k-r} f(x).$$

ainsi, nous pouvons écrire

$$\begin{aligned} & \frac{1}{k}\sum_{r=1}^{k} f_{(r)}^{2}(x) \\ = & \frac{1}{k}\sum_{r=1}^{k}\left[\frac{k!}{(r-1)!\,(k-r)!}F^{r-1}(x)\left[1-F(x)\right]^{k-r} f(x)\right]^{2} \\ = & kf^{2}(x)\sum_{j=0}^{k-1}\left[\binom{k-1}{j}F^{j}(x)\left(1-F(x)\right)^{k-1-j}\right]^{2}. \end{aligned}$$

La première partie du lemme est prouvée. La deuxième partie découle de l'exitension par Edgeworth de la probabilité $P(Y=Z)$. ■

Remarque 2.2 *Notre calcul pour certaines valeurs de* $F(x)$ *à révélé que l'approximation de la probabilité* $P(Y=Z)$ *est assez précise pour un* K *grand ou modéré. Pour un petit* k, $P(Y=Z)$ *est légèrement plus grand que l'approximation. Cependant, l'approximation peut très bien servir notre objectif théorique. Dans ce qui suit on note, pour toute fonction* g, *l'intégrale* $\int x^{\ell}g(x)\,dx$ *par* $i_{\ell}(g)$, *En appliquant les lemmes 2.1 et 2.2, nous avons*

Lemme 2.4 *Si le classement est parfait, alors pour un* k *grand ou modéré déterminé, et comme* $n\to\infty$, *nous avons*

$$MISE\left(\hat{f}_{SRS}\right)=MSIE\left(\hat{f}_{SRS}\right)-\frac{1}{n}\left[\sqrt{k}\delta(f)-i_{0}\left(f^{2}\right)\right]+O\left(\frac{h^{2}}{n}\right)$$

Lemme 2.5 *Le lemme 2.4 montre que l'estimation RSS réduit le* $MISE$ *de l'estimation SRS à l'ordre* $O(n^{-1})$ *d'un montant qui augmente de manière linéaire en* $\sqrt{k}$. *Les résultats obtenus dans cette section peuvent être étendus directement à l'estimation adaptative du noyau. L'estimation du noyau ordinaire souffre généralement d'un léger inconvénient, à savoir qu'elle a tendance à sous-aplanir à la fin de la distribution. L'estimation du noyau adaptatif surmonte cet inconvénient et fournit de meilleures estimations au niveau des queues. Nous ne discutons pas davantage de l'estimation du noyau adaptatif.*

2.5.3 Efficacité relative de l'estimation de la densité RSS par rapport à son homologue SRS

Dans cette section, nous examinons l'efficacité ralative de l'estimation RSS par rapport à l'estimation SRS en termes de ratio(taux) du $MISE$.

Tout d'abord, nous obtenons une expansion asymptotique pour l'estimation $MISE$ de SRS. Par Taylor, l'expansion de la fonction de densité f en x sous les intégrales dans (2.4) et (2.5) après avoir effectué le changement de variable $y = (x-t)/h$, nous avons

$$\begin{aligned} bias\left(\hat{f}_{SRS}\right) &= \frac{1}{2} i_2(k) f''(x) h^2 + O\left(h^4\right), \\ Var\left(\hat{f}_{SRS}(x)\right) &= \frac{1}{nh} i_0\left(k^2\right) f(x) - \frac{1}{n} f^2 + O\left(\frac{h^2}{n}\right). \end{aligned}$$

par conséquent

$$\begin{aligned} MISE\left(\hat{f}_{SRS}\right) &= \int \left[Var\left(\hat{f}_{SRS}\right) + bias^2\left(\hat{f}_{SRS}\right)\right] dx \\ &= \frac{1}{nh} i_0\left(k^2\right) + \frac{1}{4} i_2^2(k)\, i_0\left(f''^2\right) + \frac{1}{n} i_0\left(f^2\right) \\ &\quad + O\left(h^6\right) + O\left(\frac{h}{n}\right). \end{aligned} \tag{2.11}$$

En minimisant les termes principaux par rapport à h, nous avons

$$h_{opt} = i_2(k)^{-2/5} \left[\frac{i_0\left(k^2\right)}{i_0\left(f''^2\right)}\right]^{1/5} n^{-1/5}. \tag{2.12}$$

substituant (2.12) dans (2.11) les rendements

$$MISE\left(\hat{f}_{SRS}\right) = \frac{5}{2} C(k)\, i_0\left(f''^2\right)^{1/5} n^{4/5} - i_0\left(f^2\right) n^{-1} + O\left(n^{-6/5}\right), \tag{2.13}$$

où $C(k) = i_2(k)^{2/5}\, i_0\left(k^2\right)^{4/5}$.

En combinant (2.13) avec le théorèm 2.8, nous obtenons que l'efficacité relative de l'estimation RSS par rapport à l'estimation SRS est approximée par

$$\frac{MISE\left(\hat{f}_{SRS}\right)}{MISE\left(\hat{f}_{RSS}\right)} \approx \left[1 - \frac{\Delta(f,k)}{(5/4)\, C(k)\, i_0\left(f''^2\right)^{1/5} n^{1/5} - i_0\left(f^2\right)}\right]^{-1}. \tag{2.14}$$

Lorsque rang est parfait et que k est grand ou modéré, l'efficacité relative à l'expression approximative :

$$\frac{MISE\left(\hat{f}_{SRS}\right)}{MISE\left(\hat{f}_{RSS}\right)} \approx \left[1 - \frac{\sqrt{k}\delta(f) - i_0\left(f^2\right)}{(5/4)\, C(k)\, i_0\left(f''^2\right)^{1/5} n^{1/5} - i_0\left(f^2\right)}\right]^{-1} \tag{2.15}$$

Nous pouvons conclure qualitativement de l'approximation de (2.15) que

(i) l'efficacité de l'estimation du noyau RSS par rapport à l'estimation du noyau SRS augmente à mesure que k augmente au taux $O\left(k^{1/2}\right)$,

(ii) la valeur relative l'efficacité s'atténue au fur et à mesure que n augmente, mais la vitesse à laquelle il s'atténue est très faible (d'ordre $O\left(n^{-1/5}\right)$. Par conséquent, on peut s'attendre à ce que, pour les échantillons de petite taille ou de taille moyenne, le gain d'efficacité résultant de l'utilisation de RSS sera substantiel. RSS ne peut que réduire la variance et l'ordre $O\left(k^{-1}\right)$ à lequel la variance est réduite est commun dans toutes les autres procédures statistiques telles que l'estimation de la moyenne, la variance et la distribution cumulative, etc. Cependant, alors que la réduction de MISE est à l'ordre $O\left(n^{-1}\right)$, les MISEs ont l'ordre $O\left(n^{-4/5}\right)$. Lorsque n est grand, la composante du MISE à l'ordre $O\left(n^{-4/5}\right)$ est dominante. Ceci explique le fait que l'efficacité relative s'atténue au fur et à mesure que n va à l'infini.

Une application majeure de l'estimation de la densité de RSS consiste à estimer la densité à certains points particuliers, par exemple certains quantiles. Il est souhaitable de comparer les performances de l'estimation RSS et de l'estimation SRS à des valeurs particulières de x. Un argument similaire à la comparaison globale conduit aux résultats suivants.

2.6 Echantillonnage classé bivarié

En 2002 Al-Saleh & Zheng[1] ont récemment proposé une généralisation de la méthode d'échantillonnage classé de McIntyre à l'estimation de deux caractéristiques simultanées.De manière à expliquer leur approche le plus simplement possible, nous allons nous limiter ici au cas de $k = 2$ unités.

Dans cette situation, les paires de rangs possibles des deux caractéristiques sont :

$$(1;1);\ (1;2);\ (2;1);\ (2;2).$$

Instinctivement, nous pourrions penser procéder de la même manière que nous le faisions dans l'échantillonnage classé univarié,c'est-à-dire sélectionner la paire $(1;1)$ à partir d'un premier échantillon aléatoire simple, la paire $(1;2)$ d'un deuxième échantillon, et ainsi de suite. Ces paires constituant notre échantillon classé, nous pourrions alors les utiliser pour prendre les mesures nécessaires.Ces paires constituant notre échantillon classé, nous pourrions alors les utiliser pour prendre les mesures nécessaires.

Le principal écueil lié à cette approche vient du fait que la recherche de représentants de chacune des quatre paires pourrait s'avérer laborieuse

en pratique.En effet, dans le cas où les deux caractéristiques sont fortement corrélées positivement, les paires $(1;2)$ et $(2;1)$ pourraient s'avérer rares. De plus, le seul fait de juger le rang de deux caractéristiques en même temps rend cette méthode très fastidieuse.

Pour pallier ce problème,ont proposé une approche différente, qui prend appui dans le fait que l'on peut simuler une observation aléatoire d'une population bivariée en générant d'abord une observation de la loi marginale de la première variable, puis en générant une observation de la seconde à partir de sa loi conditionnelle sachant la valeur de la première variable.

Dans le cas particulier où $k = 2$, l'approche proposée par Al-Saleh & Zheng 2002 [1] consiste à prélever huit échantillons aléatoires de taille 2 de la population. Ces huit échantillons sont ensuite divisés en quatre groupes de deux échantillons chacun. La paire $(1;1)$ sera alors extraite du premier groupe, la paire $(1;2)$ sera extraite du deuxième groupe, et ainsi et suite.

Pour obtenir la paire *(1 ;1)*, on procède d'abord à un tri de chacun des deux échantillons du premier groupe et on ne retient que la paire correspondant au minimum de la première caractéristique. Ceci nous laisse donc deux paires d'observations. L'élément *(1 ;1)* de l'échantrillon bivarié est alors celle de ces deux paires pour laquelle on juge

que la deuxième caractéristique est minimale.

La procédure à suivre pour l'obtention des paires *(1 ;2),(2 ;1)* et *(2 ;2)* est semblable. Pour plus de clarté, voici, étape par étape, la procédure à suivre afin d'obtenir un échantrillon bivarié (ETB) :

1. Pour une taille k, nous avons besoin d'un échantillon de k^4 unités de la population visée.
2. Nous divisons aléatoirement ces k^4 unités en k^2 groupes de k^2 unités chacun. Chaque groupe a ainsi la forme d'une matrice carrée à k rangées et k colonnes.
3. Dans le premier groupe, nous identifions la valeur minimum de la première caractéristique pour chacune des k rangées.
4. Parmi chaque minimum ainsi obtenu, nous choisissons la paire ayant la valeur minimale de la seconde caractéristique. Nous obtenons ainsi la paire $(1, 1)$, notre premier élément de l'échantrillon bivarié.
5. Ensuite, nous répétons les étapes 3 et 4 dans le deuxième groupe, à la différence que nous prenons la paire ayant le deuxième minimum pour la deuxième caractéristique. Ceci conduit au choix de l'élément $(1, 2)$ de l'échantrillon.
6. Nous continuons ce processus jusqu'à ce que nous ayons extrait la paire $(k; k)$ du dernier groupe.

En procédant de cette manière, nous pouvons obtenir un échantrillon bivarié de taille k^2.Comme pour l'échantrillonnage univarié, même si nous utilisons seulement k^2 des k^4 unités, toutes les unités apportent de l'information sur les k^2 unités qui seront éventuellement mesurées.

2.7 Notation

Supposons que nous ayons un échantillon aléatoire de k^2 groupes carrés de taille k^2 chacun. Les éléments de chaque groupe sont supposés avoir été divisés aléatoirement en k ensembles de taille k.

Notons les valeurs des deux caractéristiques des éléments dans le n^e groupe par

$$\left\{X_{ij}^{(n)}, Y_{ij}^{(n)}, i = 1, ..., k,\ j = 1, ..., k\right\}, \quad n = 1, ..., k^2.$$

Ici $X_{ij}^{(n)}$ représente la valeur de la première caractéristique pour le j^e élément de la i^e rangée du n^e groupe. De même,$Y_{ij}^{(n)}$ représente la valeur de la deuxième caractéristique pour le j^e élément de la i^e rangée dans le n^e groupe.

Enfin, introduisons les notations supplémentaires suivantes, pour tous $i \in \{1, ..., k\}$, $j \in \{1, ..., k\}$ et $n \in \{(j-1)\,k + 1, ..., jk\}$:

$X_{i(j)}^{(n)}$ le j^e élément le plus petit parmi $X_{i1}^{(n)}$,...,$X_{ik}^{(n)}$;

$Y_{i[j]}^{(n)}$ la valeur correspondante de la variable Y ;

$Y_{(i)[j]}^{(n)}$ le i^e élément le plus petit parmi les $Y_{1[j]}^{(n)}, ..., Y_{k[j]}^{(n)}$

$X_{[i](j)}^{(n)}$ la valeur correspondante de la variable X .

Ainsi, un échantillon classé bivarié est constitué de k^2 paires

$$\left(X_{[i](j)}^{(n)}, Y_{(i)[j]}^{(n)}\right), \quad i = 1, ..., k, \quad j = 1, ..., k, \quad n = (j-1)\,k + i$$

Il est important pour la suite de garder à l'esprit que ces dernières variables sont indé pendantes, quoique non identiquement distribuées.

2.7.1 Résultats préliminaires

Nous présentons ici deux résultats qui jouent un rôle important dans la suite. La démonstration du premier est immédiate.

Lemme 2.6 *soit* $(X_1, Y_1), ..., (X_k, Y_k)$ *une échontillon aléatoire de densité* $f_{x,y}$. *Notons respectivement par* f_x *et par* $f_{y/x}$ *la densité de* X *et la densité conditionnelle de* Y *sachant* X .*Appelons* $Y_{[i]}$ *la variable concomitante de la statistique d'ordre* $X_{(i)}$ *pour tout* $i \in \{1, ...k\}$.*Alors la densité conditionnelle de* $Y_{[i]}$ *sachant que* $X_{(i)} = x$ *est donnée*

par

$$fy_{[i]}x_{(i)}\left(y|X_{(i)} = x\right), \quad i \in \{1, ..., k\}.$$

Notons par

$$fx_{i(j)}y_{i(j)} \text{ et } f_{x_{[i](j)}y_{(i)[j]}}(x, y)$$

les densités respectives des vecteurs aléatoires

$$\left(X^{(n)}_{i(j)}, Y^{(n)}_{i[j]}\right) \text{ et } \left(X^{(n)}_{[i](j)}, Y^{(n)}_{(i)[j]}\right).$$

Les densités marginales et conditionnelles sont notées de manière similaire.

Remarquons que par construction, les paires $\left(X^{(n)}_{i(j)}, Y^{(n)}_{i[j]}\right)$ sont indépendantes et de même loi peu importe les valeurs de i et de n . Dapre le lamme 2.6,on sait en outre que

$$fx_{i(j)}y_{i[j]}(x, y) = f_{x_{(j)}}(x) f_{y|x}(y|x),$$

de part et d'autre de la dernière équation, où $f_{x_{(j)}}(x)$ est la densité de la j^e statistique d'ordre d'un échantillon aléatoire de taille k de la variable X . En sommant i et sur j de de part et d'autre de la dernière équation, nous obtenons

$$\sum_{i=1}^{k}\sum_{i=1}^{k} f_{x_{i(j)}, y_{i[j]}}(x, y) = \sum_{j=1}^{k}\sum_{i=1}^{k} f_{x_{(j)}}(x) f_{y_{|x}}(y|x) = k^2 f_{x,y}(x, y).$$

Cette dernière égalité est vraie en vertu du Lemme 1.1. Par conséquent, la moyenne des densités associées aux paires $\left(X^{(n)}_{i(j)}, Y^{(n)}_{i[j]}\right)$ est donc égale à $f_{x,y.}$ De façon semblable, nous avons

$$\begin{aligned} f_{x_{[i](j)}y_{(i)(j)}}(x, y) &= f_{y_{(i)[j]}}(y) f_{x_{[i](j)}|y_{(i)[j]}}(x|y) \\ &= f_{y_{(i)[j]}}(y) f_{x_{(j)}|y_{(i)}}(x|y) \end{aligned}$$

car $X_{[i](j)}$ est concomitante de $Y_{(i)[j]}$. En appliquant le Lemme 2.6, on trouve alors

$$\begin{aligned} f_{x_{[i](j)}y_{(i)(j)}}(x,y) &= f_{y_{(i)[j]}}(y)\,\frac{f_{x_{(j)}y_{[i]}}(x,y)}{f_{y_{i[j]}}(y)} \\ &= f_{y_{(i)[j]}}(y)\,\frac{f_{x_{(j)}}(x)\,f_{y|x}(y|x)}{f_{y_{j[j]}}(y)} \end{aligned}$$

puisque $f_{y_{[j]}}(y) = f_{y_{j[j]}}(y)$
En sommant sur i de part et d'autre de l'identité, on obtient

$$\begin{aligned} \sum_{i=1}^{k} f_{x_{[i](j)}y_{(i)(j)}}(x,y). &= f_{x_{(i)}}(x)\,f_{y|x}(y|x)\sum_{i=1}^{k}\frac{f_{y_{(i)[j]}}(y)}{f_{y_{j[j]}}(y)} \\ &= kfx_{[i](j)}(x)\,f_{y|x}(y|x), \end{aligned}$$

par application du Lemme 1.1.
En intégrant par rapport à y de chaque ctôét, on voit alors que

$$\sum_{i=1}^{k} f_{x_{[i](j)}}(x) = kf_{x_{(j)}}(x)$$

Si au lieu d'intégrer on somme plutôt par rapport à j, une nouvelle invocation du Lemme 1.1 permet de conclure que

$$\sum_{i=1}^{k}\sum_{i=1}^{k} f_{y_{i[j]}}(x,y) = kf_{y|x}(y|x)\sum_{j=1}^{k} f_{x_{(j)}}(x) = k^2 f_{x,y}(x,y).$$

Ces faits sont énoncés formellement ci-dessous, de même qu'une de leurs conséquences immédiates.

Lemme 2.7 *Soit (X,Y) une paire de variables aléatoires ayant pour densité $f_{x,y}$. Pour tout $i \in \{1,...,k\}$, $j = \{1,...,k\}$ et $n = (j-1)k + i$, soit $fx_{[i](j)y_{(i)[j]}}$ la dénsité de $\left(X^{(n)}_{[i](j)}, Y^{(n)}_{(i)[j]}\right)$. Alors*

$$f_{x,y}(x,y) = \frac{1}{k^2}\sum_{i=1}^{k}\sum_{i=1}^{k} fx_{[i](j)y_{(i)[j]}}(x,y)$$

De plus, $f_{x_{(j)}}(x) = \frac{1}{k}\sum_{i=1}^{k}\sum_{i=1}^{k} f_{x_{[i](j)}}(x)$ et $f_x(x) = \frac{1}{k^2}\sum_{i=1}^{k}\sum_{i=1}^{k} fx_{[i](j)}(x)$.

Remarque 2.3

(i) *si X et Y sont indépendants, alors* $\left(X_{[i](j)}^{(n)}, Y_{(i)[j]}^{(n)}\right)$ *et* $\left(X_{i(j)}^{(n)}, Y_{(i)j}^{(n)}\right)$ *sont de même loi et les valeurs de* $X_{i(j)}^{(n)}$ *et de* $Y_{(i)j}^{(n)}$ *constituent des échantrillons univariés à n cycles de taille k respectivement construits à partir des densités f_x et f_y. Dans ce cas, l'échantrillonnage bivarié est équivalent à deux échantrillons univariés de taille k^2 chacun.*

(ii) *Si X et Y sont parfaitement corrélés, de sorte que $p(X,Y) = 1$,alors*

$$\left(X_{[i](j)}^{(n)}, Y_{(i)[j]}^{(n)}\right) \text{ et } \left(X_{i(j)}^{(n)}, Y_{(i)j}^{(n)}\right)$$

sont de même loi, de sorte que

$$\left\{X_{(i)(j)}^{(n)},\ i = 1, ..., k,\right\} \text{ et } \left\{Y_{(i)(j)}^{(n)},\ j = 1, ..., k\right\}$$

se réduisent à des échantrillons univariés à un cycle de taille k respectivement obtenus à partir de $f_{X_{(i)}}$ et $f_{Y_{(I)}}$, où $i \in \{1, ..., k,\}$ et $n = (j-1)k + i$.

2.7.2 Les moyennes échantillonnales classé

Dénotons par

$$\left\{\left(X_{[i](j)}^{(n)}, \left\{Y_{(i)[j]}^{(n)}\right\}\right),\ i = 1, ..., k,\ j = 1, ..., k,\ n = (j-1)k + i\right\}$$

un échantillon classé bivarié de taille k^2 prélevé à partir d'une population bivariée ayant pour densité $f_{X,Y}$. On suppose comme précédemment que les rangs ne sont sujets à aucune erreur de jugement. Dénotons respectivement par μ et par θ l'espérance et la variance de X. De la même manière, soient θ et σ_Y^2 l'espérance et la variance de Y. Enfin, soit le coefficient de corrélation entre X et Y.

Supposons que l'on veuille estimer μ et θ. Appelons $\hat{\mu}_{ETB}$ la moyenne échantillonnale classé de

$$\left\{X_{[i](j)}^{(n)},\ i = 1, ..., k,\ j = 1, ..., k\right\}$$

et $\hat{\theta}_{ETB}$ la moyenne échantillonnale classé de

$$\left\{Y_{(i)[j]}^{(n)}, i = 1, ..., k,\ j = 1, ..., k\right\}.$$

Posons aussi

$$\mu_{[i](j)} = E\left(X_{[i](j)}^{(n)}\right),\ \theta_{(i)[j]} = E\left(Y_{[i](j)}^{(n)}\right)$$

et

$$\sigma^2_{(i)[j]} = Var\left(X^{(n)}_{[i](j)}\right).$$

En vertu du Lemme *2.7*, nous avons

$$\frac{1}{k^2}\sum_{j=1}^{k}\sum_{i=1}^{k} E\left(Y^{(n)}_{[i](j)}\right) = \frac{1}{k^2}\sum_{j=1}^{k}\sum_{i=1}^{k}\theta^{(n)}_{(i)[j]} = \theta.$$

Ainsi, $\hat{\theta}_{ETB}$ et $\hat{\mu}_{ETB}$ sont respectivement des estimateurs sans biais de μ et de θ, comme nous l'avons déjà montré au chapitre *1.5* dans le cas de l'échantrillonnage univarié.

2.7.3 La variance des moyennes échantillonnales classé

Deux expressions différentes seront données ci-dessous pour le calcul de la variance de l'estimateur $\hat{\mu}_{ETB}$ de $\mu = E(X)$. Des formules semblables peuvent être obtenues de façon semblable pour la variance de la moyenne échantrillonnale $\hat{\theta}_{ETB}$.

Notons d'abord que par une application du Lemme 2.7, on a

$$\begin{aligned}
\sigma^2_X &= \int_{-\infty}^{\infty}(x-\mu)^2 f_X(x)\,dx = \frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k}\int_{-\infty}^{\infty}(x-\mu)^2 f_{X_{[i](j)}}(x)\,dx \\
&= \frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k}\int_{-\infty}^{\infty}\left(x-\mu_{[i](j)}+\mu_{[i](j)}-\mu\right)^2 f_{X_{[i](j)}}(x)\,dx \\
&= \frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k}\int_{-\infty}^{\infty}\left(x-\mu_{[i](j)}\right)^2 f_{X_{[i](j)}}(x)\,dx \\
&\quad + \frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k}\int_{-\infty}^{\infty}\left(\mu_{[i](j)}-\mu\right)^2 f_{X_{[i](j)}}(x)\,dx,
\end{aligned}$$

puisque le terme croisé s'annule du fait que par définition,

$$\int\left(x-\mu_{[i](j)}\right)^2 f_{X_{[i](j)}}(x)\,dx = 0$$

pour tous $i, j \in \{1, ..., k\}$. Il s'ensuit que

$$\sigma^2_X = \frac{1}{k^2}\sum_{j=1}^{k}\sum_{i=1}^{k}\sigma^2_{[i](j)} + \frac{1}{k^2}\sum_{j=1}^{k}\sum_{i=1}^{k}\left(\mu_{[i](j)}-\mu\right)^2$$

et que par conséquent,

$$\begin{aligned} Var\left(\hat{\mu}_{ETB}\right) &= Var\left(\frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k}X_{[i](j)}^{(n)}\right) = \frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\sigma_{[i](j)}^2 \\ &= \frac{\sigma_X^2}{k^2} - \frac{1}{k^4}\sum_{j=1}^{k}\sum_{i=1}^{k}\left(\mu_{[i](j)} - \mu\right)^2. \end{aligned} \tag{2.16}$$

Une deuxiéme expression pour la variance de $\hat{\mu}_{ETB}$ fait intervenir l'espérance et la variance des variables $X_{i(j)}$, notées

$$\mu_j = E\left(X_{i(j)}\right), \quad \sigma_j^2 = Var\left(X_{i(j)}\right)$$

pour tous $i, j \in \{1, ..., k\}$. Cette formule, semblable à celle déjà donnée dans le cas univarié, s'obtient en deux temps.

Notons d'abord que

$$\begin{aligned} Var\left(\hat{\mu}_{ETB}\right) &= \frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}Var\left(X_{[i](j)}^{(n)}\right) \\ &= \frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\int\left(x-\mu_{[i](j)}\right)^2 f_{X_{[i](j)}}(x)\,dx \\ &= \frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\int\left(x-\mu_j+\mu_j-\mu_{[i](j)}\right)^2 f_{X_{[i](j)}}(x)\,dx \\ &= \frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\int\left(x-\mu_j\right)^2 f_{X_{[i](j)}}(x)\,dx \\ &\quad + \frac{2}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\mu_j-\mu_{[i](j)}\right)\int\left(x-\mu_j\right) f_{X[i](j)}(x)\,dx \\ &\quad + \frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\mu_j-\mu_{[i](j)}\right)^2\int f_{X[i](j)}(x)\,dx \end{aligned}$$

et donc que

$$\begin{aligned} Var\left(\hat{\mu}_{ETB}\right) &= \frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\int\left(x-\mu_j\right)^2 f_{X_{[i](j)}}(x)\,dx \\ &\quad - \frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\mu_j-\mu_{[i](j)}\right)^2. \end{aligned} \tag{2.17}$$

Or, d'après le Lemme 2.7, on sait que

$$\sum_{i=1}^{k} f_{X(j)} = k f_{X_{(j)}}.$$

Par substitution dans le premier terme du membre de droite de la formule *(2.17)*, on trouve alors

$$Var\left(\hat{\mu}_{ETB}\right) = \frac{1}{k^3}\sum_{i=1}^{k}\int \left(x-\mu_j\right)^2 f_{X_{(j)}}\left(x\right)dx - \frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\mu_j - \mu_{[i](j)}\right)^2.$$

2.7.4 Comparaisons élémentaires

Pour juger de l'efficacité de l'estimation de $\mu = E\left(X\right)$ par échantrillonnage bivarié, on peut comparer la variance de $\hat{\mu}_{ETB}$ à deux autres estimations fondées sur des efforts de mesure semblables, à savoir :

a) l'estimateur $\hat{\mu}_{EAS}$ fondé sur un échantillon aléatoire simple de taille k^2 ;

b) l'estimateur $\hat{\mu}_{ETU}$ fondé sur un échantrillon à k cycles de taille k chacun.

Les variances de ces deux compétiteurs sont les suivantes :

$$\begin{aligned} Var\left(\hat{\mu}_{EAS}\right) &= \sigma_X^2/k^2, \\ Var\left(\hat{\mu}_{ETU}\right) &= \frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k} Var\left(X_{i(j)}\right) = \frac{1}{k^3}\sum_{j=1}^{k}\sigma_j^2. \end{aligned} \quad (2.18)$$

Au vu de l'équation *(2.16)*, l'efficacité relative de $\hat{\mu}_{ETB}$ par rapport à $\hat{\mu}_{EAS}$ est donc donnée par

$$eff\left(\hat{\mu}_{ETB}|\hat{\mu}_{EAS}\right) = \frac{Var\left(\hat{\mu}_{EAS}\right)}{Var\left(\hat{\mu}_{ETB}\right)} = \frac{\sigma_X^2/k^2}{\frac{\sigma_X^2}{k^2} - \frac{1}{k^4}\sum_{j=1}^{k}\sum_{i=1}^{k}\left(\mu_{[i](j)}-\mu\right)^2}. \quad (2.19)$$

Le rapport des variances étant supérieur à 1, on voit qu'en général, $\hat{\mu}_{EAS}$ est plus efficace que $\hat{\mu}_{ETB}$ pour un même effort d'échantillonnage, soit k^2. Il en va évidemment de même pour $\hat{\theta}_{ETB}$ par rapport à $\hat{\theta}_{EAS}$.

En faisant appel μa la formule *(2.18)*, on voit en outre que

$$eff\left(\hat{\mu}_{ETB}|\hat{\mu}_{EAS}\right) = \frac{Var\left(\hat{\mu}_{ETU}\right)}{Var\left(\hat{\mu}_{ETB}\right)} = \frac{\frac{1}{k^3}\sum_{j=1}^{k}\sigma_j^2}{\frac{1}{k^3}\sum_{j=1}^{k}\sigma_j^2 - \frac{1}{k^4}\sum_{j=1}^{k}\sum_{i=1}^{k}\left(\mu_{[i](j)}-\mu_j\right)^2}. \quad (2.20)$$

Une fois de plus, le rapport est donc supérieur à 1, ce qui montre qu'une estimation de μ fondée sur un échantrillon bivarié est préférable à une estimation fondée sur un échantrillon univarié à k cycles.

2.7.5 Comparaisons plus fines

Nous allons maintenant affiner nos comparaisons d'efficacité entre les estimateurs $\hat{\mu}_{ETB}$ et $\hat{\mu}_{ETU}$ en considérant trois scénarios de dépendance particuliers. Avant de procéder, remarquons qu'en vertu de l'identité *(1.1)*, on a

$$Var\left(\hat{\mu}_{ETU}\right) = \frac{1}{k^3}\sum_{j=1}^{k}\sigma_j^2 = \frac{\sigma_X^2}{k^2} - \frac{1}{k^3}\sum_{j=1}^{k}\left(\mu_j - \mu\right)^2.$$

Etant donné l'identité *(2.18)*, on a donc aussi

$$Var\left(\hat{\mu}_{ETB}\right) = \frac{\sigma_X^2}{k^2} - \frac{1}{k^3}\sum_{j=1}^{k}\left(\mu_j - \mu\right)^2 - \frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\mu_{[i](j)} - \mu_j\right). \quad (2.21)$$

Le cas d'indépendance

Si les variables X et Y sont indépendantes, alors $\mu_{[i](j)} = \mu_j$ pour tout $i \in \{1, ..., k\}$.

De toute évidence, le troisiμeme membre du terme de droite de l'équation *(2.21)* est alors identiquement égal à zéro.

Par conséquent, on a

$$eff_0\left(\hat{\mu}_{ETB}|\hat{\mu}_{EAS}\right) = 1,$$

ce qui signifie que l'échantrillonnage bivarié est aussi efficace que l'échantrillonnage univarié lorsque X et Y sont indépendants. Sous ces conditions, l'échantrillonnage bivarié peut tout de même s'avérer plus avantageux que l'approche univariée dans certaines circonstances. Cette façon de procéder pourrait être plus économique, par exemple, lorsque la prise de mesure revêt un caractère destructeur.

Le cas de dépendance parfaite

Si les variables X et Y sont parfaitement corrélées positivement ($p = 1$), alors

$$\left(X_{[i](j)}^{(n)}, Y_{(i)[j]}^{(n)}\right) \text{ et } \left(X_{(i)(j)}^{(n)}, Y_{(i)(j)}^{(n)}\right)$$

sont de même loi, de sorte que $\mu_{[i](j)} = \mu_{(i)(j)}$ pour tous $i, j \in \{1, ..., k\}$. Il découle alors de l'équation (2.20) que

$$eff_1\left(\hat{\mu}_{ETB}|\hat{\mu}_{EAU}\right) = \frac{\frac{1}{k^3}\sum_{j=1}^{k}\sigma_j^2}{\frac{1}{k^3}\sum_{j=1}^{k}\sigma_j^2 - \frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\mu_{(i)(j)} - \mu_j\right)^2} \quad (2.22)$$

est supérieur à 1.

Le cas de dépendance linéaire

A l'instar de Stokes (1977)[22], qui fait cette hypothèse dans le cas univarié, supposons de façon plus générale que la régression de X en Y soit linéaire, comme ce serait le cas par exemple si la paire (X, Y) obéissait à une loi normale ou à une loi de Pareto bivariée. Cette relation de linéarité n'étant pas affectée par des opérations de tri et de classement, on a alors

$$\mu_{[i](j)} = \mu_j + p\left(\mu_{(i)[j]} - \mu_j\right)$$

pour un certain coefficient de corrélation $p \in [-1, 1]$. Noter qu'en particulier on retrouve les relations $\mu_{[i](j)} = \mu_j$ et $\mu_{[i](j)} = \mu_{(i)(j)}$ correspondant respectivement à l'indépendance $(p = 0)$ et à la dépendance positive parfaite $(p = 1)$.

Sous ce modèle particulier de dépendance, l'équation *(2.20)* devient

$$eff_p\left(\hat{\mu}_{ETB}|\hat{\mu}_{EAU}\right) = \frac{\frac{1}{k^3}\sum_{j=1}^{k}\sigma_j^2}{\frac{1}{k^3}\sum_{j=1}^{k}\sigma_j^2 - \frac{p^2}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\mu_{(i)(j)} - \mu_j\right)^2}. \qquad (2.23)$$

Or à la lumière de l'identité *(2.23)*, on sait que

$$\frac{\frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\mu_{(i)(j)} - \mu_j\right)^2}{\frac{1}{k^3}\sum_{j=1}^{k}\sigma_j^2} = 1 - \frac{1}{eff_1\left(\hat{\mu}_{ETB}|\hat{\mu}_{EAU}\right)}.$$

En exploitant ce fait, il est donc possible d'exprimer la formule d'efficacité générale comme suit, en fonction du paramètre p et de l'efficacité en $p = 1$:

$$\Phi(p) = eff_p\left(\hat{\mu}_{ETB}|\hat{\mu}_{EAU}\right) = \frac{eff_1\left(\hat{\mu}_{ETB}|\hat{\mu}_{EAU}\right)}{eff_1\left(\hat{\mu}_{ETB}|\hat{\mu}_{EAU}\right) - p^2\left\{eff_1\left(\hat{\mu}_{ETB}|\hat{\mu}_{EAU}\right) - 1\right\}}. \qquad (2.24)$$

De plus, puisque la même équation *(2.23)* traîne que

$$eff_1\left(\hat{\mu}_{ETB}|\hat{\mu}_{EAU}\right) \geqslant 1,$$

on déduit facilement de *(2.24)* que $\Phi(p)$ est une fonction croissante de $|p|$ dont le minimum est atteint en $\Phi(0) = 1$, en accord avec le résultat déjà énoncé dans la sous-section Un exemple de calcul explicite de la fonction sera $\Phi(p)$ présenté ci-dessous.

2.7.6 Cas de la loi normale bivariée

Afin d'illustrer l'efficacité de l'estimation par échantillonnage classé bivarié, nous nous penchons dans cette section sur le cas où la population est normale bivariée, c'est-à-dire où

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}_2 \left[\begin{pmatrix} \mu \\ \theta \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & p\sigma_X\sigma_Y \\ p\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix} \right].$$

Comme précédemment, concentrons-nous sur l'estimation de μ et comparons la performance de l'estimateur $\hat{\mu}_{ETB}$ à celle de l'estimateur $\hat{\mu}_{ETS}$ correspondant à un échantillon aléatoire de taille k^2.

Partons du fait que

$$\begin{aligned} eff_p\left(\hat{\mu}_{ETB}|\hat{\mu}_{EAS}\right) &= eff_p\left(\hat{\mu}_{ETB}|\hat{\mu}_{EAU}\right) * eff\left(\hat{\mu}_{ETU}|\hat{\mu}_{EAS}\right) \\ &= \Phi\left(p\right) * eff\left(\hat{\mu}_{ETU}|\hat{\mu}_{EAS}\right). \end{aligned}$$

Puisque le modéle normal répond aux hypothèses ayant conduit à l'identité *(2.24)*, on sait que $1/\Phi\left(p\right)$ est une fonction quadratique de p, dont le comportement dépend exclusivement de $\Phi\left(1\right)$. L'efficacité relative de l'estimateur $\hat{\mu}_{ETB}$ ne dépend donc que de cette constante et de $eff\left(\hat{\mu}_{ETU}|\hat{\mu}_{EAS}\right)$.

Or, compte tenu de la formule *(1.1)*, on sait déjà que

$$eff\left(\hat{\mu}_{ETU}|\hat{\mu}_{EAS}\right) = \frac{\sigma_X^2/k}{\frac{\sigma_X^2}{k} - \frac{1}{k^2}\sum_{j=1}^{k}\left(\mu_j - \mu\right)^2} = \frac{1}{1 - \frac{1}{k}\sum_{j=1}^{k}\zeta_j^2},$$

où pour tout $j \in \{1, ..., k\}$,

$$\zeta_{(j)} = E\left(\frac{X_{(j)} - \mu}{\sigma_X}\right) = \frac{\mu_j - \mu}{\sigma_X}$$

est l'espérance de la je statistique d'ordre d'une variable normale centrée réduite, aussi appelée je rankit. Ces constantes, indépendantes de μ et de σ_X, sont tabulées dans le livre de David (1981)[9] et, de toute façon, très faciles à déterminer par intégration numérique. Il ne reste donc qu'à déterminer la valeur de $eff_1\left(\hat{\mu}_{ETB}|\hat{\mu}_{EAU}\right)$ dans le cas normal.

Avant de ce faire, notons qu'en général,

$$\begin{aligned} \sum_{i=1}^{k}\sum_{j=1}^{k}\left(\mu_{[i](j)} - \mu_j\right)^2 &= \sum_{i=1}^{k}\sum_{j=1}^{k}\left(\mu_{[i](j)} - \mu + \mu - \mu_j\right)^2 \\ &= \sum_{i=1}^{k}\sum_{j=1}^{k}\left\{\left(\mu_{[i](j)} - \mu_j\right)^2 + \left(\mu - \mu_j\right)^2 + 2\left(\mu_{[i](j)} - \mu\right)\left(\mu - \mu_j\right)\right\}, \\ &= \sum_{i=1}^{k}\sum_{j=1}^{k}\left(\mu_{[i](j)} - \mu_j\right)^2 - k\sum_{j=1}^{k}\left(\mu - \mu_j\right), \end{aligned}$$

puisque en vertu du Lemme 2.7 a

$$\sum_{i=1}^{k} \mu_{[i](j)} = \sum_{i=1}^{k} \int x f_{X_{[i](j)}}(x) = k \int x f_{X_{(j)}} dx = k\mu_j$$

et donc pour tout $j \in \{1, ..., k\}$,

$$\sum_{i=1}^{k} \left(\mu_{[i](j)} - \mu\right) = k\left(\mu_j - \mu\right).$$

Par suite, une formule équivalente à *(2.21)* est donnée par

$$\begin{aligned} Var\left(\hat{\mu}_{ETB}\right) &= \frac{\sigma_X^2}{k^2} - \frac{1}{k^3}\sum_{j=1}^{k}\left(\mu_j - \mu\right)^2 - \frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\mu_{[i](j)} - \mu_j\right)^2 \\ &= \frac{\sigma_X^2}{k^2} - \frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\mu_{[i](j)} - \mu\right)^2, \end{aligned}$$

alors que

$$Var\left(\hat{\mu}_{ETU}\right) = \frac{1}{k^3}\sum_{j=1}^{k}\sigma_j^2 = \frac{\sigma_X^2}{k^2} - \frac{1}{k^3}\sum_{j=1}^{k}\left(\mu_j - \mu\right)^2.$$

On conclut donc que

$$\begin{aligned} eff_1\left(\hat{\mu}_{ETB}|\hat{\mu}_{EAU}\right) &= \frac{\frac{\sigma_X^2}{k^2} - \frac{1}{k^3}\sum_{j=1}^{k}\left(\mu_j - \mu\right)^2}{\frac{\sigma_X^2}{k^2} - \frac{1}{k^4}\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\mu_{[i](j)} - \mu\right)^2} \\ &= \frac{1 - \frac{1}{k}\sum_{j=1}^{k}\zeta_j^2}{1 - \frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k}\zeta_{[i](j)}^2}, \end{aligned}$$

où par définition,

$$\zeta_{[i](j)} = \frac{\mu_{[i](j)} - \mu}{\sigma_X^2}, \quad i, j \in \{1, ..., k\}.$$

Ces constantes sont, elles aussi, tabulées dans le manuel de David(1981)[9]. Tout est donc en place pour donner un exemple de calcul.

Nous proposons un estimateur RSS de type régression lorsque la moyenne de la population μ_X de X est connue et aprés nous considérons ses performances et comparez-le avec d'autres estimateurs. Pour suite un double échantillonnage est utilisé pour estimerμ_X est supposé être inconnu. Comparaisons numériques de la précision relative de divers estimateurs

lorsque X et Y suivent conjointement une distribution normale à deux variables sont abordés .

2.7.7 Estimateur de régression quand μ_X est connu

Dans cette section, nous suivons l'idée de Stokes selon laquelle le classement est effectué au moyen d'une variable concomitante X facile à mesurer et corrélée positivement à la variable d'intérêt Y (coûteuse ou difficile à mesurer).D'après Stokes (1977)[22], nous supposons que la régression de Y sur X est linéaire, c-à-d.

$$Y = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(X - \mu_X) + \varepsilon_i, \tag{2.25}$$

où X et ε sont indépendants et ε a une moyenne nulle. Il en résulte que ε a une variance donnée par ou $\sigma_Y^2(1-\rho^2)$.

Soit $X_{(i)r}$ et $Y_{[i]r}$, respectivement, le $i^{ème}$ plus petit X et la valeur correspondante de Y obtenue à partir du $i^{ème}$ échantillon du $i^{ème}$ cycle. Ensuite, nous avons, de 2.25,

$$Y_{[i]r} = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}\left(X_{(i)r} - \mu_X\right) + \varepsilon_{(i)r} \quad i = 1, 2, ..., n \;\; r = 1, 2, ..., m, \tag{2.26}$$

Lorsque la moyenne μ_X de X de la population est connue, on considère l'estimateur par différence

$$\hat{Y}_D = \hat{Y}_{RSS} + B\left(\mu_X - \hat{X}_{RSS}\right),$$

où

$$\bar{Y}_{RSS} = \left(\frac{1}{mn}\right)\sum_{r=1}^{m}\sum_{r=1}^{n} Y_{[i]r},$$

$$\bar{X}_{RSS} = \left(\frac{1}{mn}\right)\sum_{r=1}^{m}\sum_{r=1}^{n} X_{(i)r},$$

et B est une constante à déterminer. $\bar{Y}_{RSS}$ et $\bar{X}_{RSS}$ étant non biaisés, on peut facilement voir que $\bar{Y}_D$ est non biaisé et, si (2.26) est satisfaite,

$$Var\left(\bar{Y}_D\right) = B^2 Var\left(\bar{X}_{RSS}\right) - 2Var\left(\bar{X}_{RSS}\right) + Var\left(\bar{Y}_{RSS}\right)$$

comme

$$\begin{aligned} Var\left(\bar{X}_{RSS}\right) &= \frac{\sigma_X^2}{mn}\left[1 - \frac{1}{n}\sum_{i=1}^{n}\left(\frac{E\left(X_{(i)r}\right) - \mu_X}{\sigma_X}\right)^2\right] \\ Var\left(\bar{Y}_{RSS}\right) &= \frac{\sigma_Y^2}{mn}\left[1 - \frac{\rho^2}{n}\sum_{i=1}^{n}\left(\frac{E\left(X_{(i)r}\right) - \mu_X}{\sigma_X}\right)^2\right] \end{aligned}$$

Comme $\bar{Y}_D$ est non biaisé pour toute valeur de B, la valeur optimale de B peut être obtenue en minimisant la variance de $\bar{Y}_D$ et est donnée par

$$B^* = \rho \frac{\sigma_Y}{\sigma_X}.$$

Cependant, B^* est inconnu et un estimateur naturel pour B^* est

$$\hat{B} = \frac{\sum_{r=1}^{m} \sum_{r=1}^{n} \left(X_{(i)r} - \bar{X}_{RSS}\right) \left(Y_{[i]r} - \bar{Y}_{RSS}\right)}{\sum_{r=1}^{m} \sum_{r=1}^{n} \left(X_{(i)r} - \bar{X}_{RSS}\right)};$$

Par conséquent, nous définissons l'estimateur de régression RSS pour μ_Y, comme suit :

$$\hat{Y}_{reg} = \hat{Y}_{RSS} + \hat{B} \left(\mu_X - \hat{X}_{RSS}\right). \tag{2.27}$$

En utilisant les propriétés de base sur les moments conditionnels, nous avons le théorème suivant. La preuve est omise ici.

Théorème 2.9 *Supposons que (2.26) soit satisfaite. Alors, l'estimateur de régression pour μ_Y défini dans (2.27) a les propriétés suivantes :*

a) $E\left(\bar{Y}_{reg}\right) = \mu_Y$,

b) $Var\left(\bar{Y}_{reg}\right) = \frac{\sigma_Y^2}{mn}\left(1-\rho^2\right)\left[1 + E\left(\frac{\bar{Z}_{RSS}}{S_Z^2}\right)\right]$,

avec

$$\begin{aligned} Z_{(i)r} &= \frac{X_{(i)r} - \mu_X}{\sigma_X}, \\ \bar{Z}_{RSS} &= \frac{1}{mn} \sum_{r=1}^{m} \sum_{r=1}^{n} Z_{(i)r} \end{aligned}$$

et

$$S_Z^2 = \frac{1}{mn} \sum_{r=1}^{m} \sum_{r=1}^{n} \left(Z_{(i)r} - \bar{Z}_{RSS}\right).$$

On peut voir que $\bar{Y}_{reg}$ est toujours non biaisé quelles que soient les distributions de X et Y. Cependant, la variance implique le terme $E\left(\frac{\bar{Z}_{RSS}}{S_Z^2}\right)$, qui dépend du choix de la distribution de $Z = \frac{X-\mu_X}{\sigma_X}$. Le calcul de ce terme lorsque X est normalement distribué.

2.7.8 Comparaison avec d'autres estimateurs

Nous pouvons déduire la précision relative de l'estimateur par régression RSS par rapport à l'estimateur naïf par RSS $\bar{Y}_{RSS}$ donnée par

$$RP\left(\bar{Y}_{reg}, \bar{Y}_{RSS}\right) = \frac{1-\rho^2 E\left(\frac{1}{n}\sum_{i=1}^{n} Z_{(i)r}^2\right)}{(1-\rho^2)\left[1+E\left(\frac{Z_{RSS}^2}{S^2}\right)\right]} \tag{2.28}$$

Notez que $\bar{Y}_{RSS}$ n'a utilisé aucune information sur la variable concomitante X. Une comparaison plus juste consiste donc à comparer l'estimateur de régression RSS avec l'estimateur de régression $\bar{Y}_{reg}$ basé sur un échantillon aléatoire simple défini par

$$\bar{Y}_{reg} = \bar{y} + \hat{\beta}\left(\mu_X - \bar{x}\right),$$

où $\bar{x}$, $\bar{y}$ et, désignent respectivement la moyenne d'échantillon de X, la moyenne d'échantillon de Y basée sur les observations en mn et l'estimation par les moindres carrés du coefficient de régression de la population β, respectivement. [Voir Hedayat et Sinha (1992)[11] pour une analyse détaillée de cet estimateur par régression]

En général, l'estimateur de régression $\bar{Y}_{reg}$ est biaisé et sa précision peut être mesurée par $MSE\left(\bar{Y}_{reg}\right)$. Mais lorsque l'échantillon est tiré d'une population normale bivariée, l'estimateur de régression $\bar{Y}_{reg}$ est connu pour être non biaisé (voir Tikkiwal, 1960)[25], et nous pouvons utiliser le ratio de variance.

$$RP\left(\bar{Y}_{reg}, \hat{Y}_{reg}\right) = \frac{Var\left(\hat{Y}_{reg}\right)}{Var\left(\bar{Y}_{reg}\right)} \tag{2.29}$$

comme mesure de la précision relative. Sukhatme et Sukhatme (1970)[23] ont montré que la variance de l'estimateur de régression $\hat{Y}_{reg}$ basée sur le SRS avec la taille de l'échantillon nm est donnée par

$$Var\left(\hat{Y}_{reg}\right) = \frac{\sigma_Y^2}{mn}\left(1-\rho^2\right)\left(1+\frac{1}{mn-3}\right), \tag{2.30}$$

à condition que l'échantillonnage soit effectué à partir d'une distribution normale bivariée.donc, la précision relative de $\bar{Y}_{reg}$ basée sur RSS par rapport à $\hat{Y}_{reg}$ basée sur SRS est donnée comme suite

$$RP\left(\bar{Y}_{reg}, \hat{Y}_{reg}\right) = \frac{1+\frac{1}{mn-3}}{1+E\left(E\left(\frac{Z_{RSS}^2}{S^2}\right)\right)}. \tag{2.31}$$

2.7.9 Le cas du double échantillonnage

Les estimateurs de régression $\bar{Y}_{reg}$ et $\hat{Y}_{reg}$ impliquent la moyenne de population μ_X de la variable concomitante X, ce qui est généralement inconnu en pratique. Si μ_X est inconnu, la méthode du double échantillonnage (ou échantillonnage à deux phases) peut être utilisée pour obtenir une estimation de μ_X. Cela implique le tirage d'un grand échantillon aléatoire de taille n', utilisé pour estimer les μ_X ; un sous-échantillon de taille n'' est sélectionné parmi les unités d'origine pour étudier la caractéristique principale Y. Sous un paramètre d'échantillonnage d'ensemble classé, $n' = n^2m$ et $n'' = nm$. Note que l'échantillonnage de la première phase est un échantillonnage aléatoire simple et que l'échantillonnage de la deuxième phase est soit un échantillonnage aléatoire simple, soit un échantillonnage classé.

Soit $\bar{x}$ la moyenne empirique de l'échantillon X basée sur des observations n^2m de X dans la première phase. Clairement, $\bar{x}$ est un estimateur sans biais pour x. Si le SRS est l'échantillonnage de la deuxième phase, l'estimateur de régression à double échantillonnage de la moyenne de la population, μ_Y est défini comme suit :

$$\hat{Y}_{ds} = \bar{y} + \hat{\beta}\left(\bar{x}' - \bar{x}\right),$$

alors que, si RSS est l'échantillonnage de deuxiém phase, il est défini comme

$$\bar{Y}_{ds} = \bar{Y}_{RSS} + \hat{B}\left(\bar{x}' - \bar{X}_{RSS}\right). \tag{2.32}$$

Nous pouvons montrer que $\bar{Y}_{ds}$ est impartial et, si (2.26) est satisfait,

$$Var\left(\bar{Y}_{ds}\right) = \frac{\sigma_Y^2\left(1-\rho^2\right)}{mn}\left[1 + E\left(\frac{\left(\bar{Z}_{RSS} - \bar{Z}\right)^2}{S_Z^2}\right) + \frac{1}{n^2m}\rho^2\sigma_Y^2,\right]$$

où $\bar{Z} = \left(\bar{x}' - \mu_X\right)/\sigma_X$.

Pour l'estimateur par régression à double échantillonnage $\hat{Y}_{ds}$, Sukhatme (1970)[23] ont montré que, si (X, Y) suit une distribution normale à deux variables, alors $\hat{Y}_{ds}$ est un estimateur sans biais de μ_Y et sa variance est donnée par

$$Var\left(\bar{Y}_{ds}\right) = \frac{\sigma_Y^2}{mn}\left(1-\rho^2\right)\left[1 + \frac{n-1}{n}\frac{1}{mn-3}\right] + \frac{1}{n^2m}\rho^2\sigma_Y^2.$$

Les analogues à double échantillonnage des équations (2.28) et (2.31)

$$RP\left(\bar{Y}_{ds}, \bar{Y}_{reg}\right) = \frac{1 - \rho^2 E\left(\frac{1}{n}\sum_{i=1}^{n} Z_{(i)r}^2\right)}{\left(1-\rho^2\right)\left[1 + E\left(\frac{\left(\bar{Z}_{RSS}-\bar{Z}\right)^2}{S_Z^2}\right)\right] + \frac{\rho^2}{n}}, \tag{2.33}$$

$$RP\left(\bar{Y}_{ds},\hat{Y}_{reg}\right)=\frac{(1-\rho^2)\left[1+\frac{n-1}{n}\frac{1}{mn-3}\right]+\frac{\rho^2}{n}}{(1-\rho^2)\left[1+E\left(\frac{\left(\bar{Z}_{RSS}-\bar{Z}\right)^2}{S_Z^2}\right)\right]+\frac{\rho^2}{n}}. \tag{2.34}$$

Chapitre 3

Applications et simulation par la méthode RSS

3.1 Introduction

L'echantillonnage par ensemble classé(RSS) utilise une information auxiliaire peu couteuse pour la determination du rang des unites d'un echantillon afin de fournir un estimateur plus précis de la moyenne de la population de la variable d'interet Y, qui est soit difficile à mesurer.Cependant, la determination du rang peut ne pas etre parfaite dans la plupart des situations.Dans cet article, nous supposons que la determination du rang est faite sur la base d'une variable concomitante X. Des estimateurs RSS de type regression sont proposes pour la moyenne de la population de Y, par l'utilisation de cette variable concomitante X, a la fois lors de la procedure de determination des rangs des unites et dans la procedure d'estimation, lorsque la moyenne de la population de X est connue,lorsque X a une moyenne inconnue,l 'echantillonnage double sera utilise pour obtenir une estimation de la moyenne de la population de X. On a trouve que lorsque X et Y suivent conjointement une distribution normale a deux dimensions, notre estimateur RSS de regression est plus efficace que les estimateurs naifs RSS et SRS, a moins que la correlation de X et Y ne soit faible ($|\rho| < 0.4$). De plus, il est toujours superieur a l'estimateur de regression dans le cas d'echantillonnage aléatoire simple (SRS) pour tout ρ. Lorsque la normalité n'est pas vérifiée, cette approche pourrait encore donner de bons resultats tant que la forme de la distribution de la variable concomitante X s'ecarte legerement de la symetrie. Pour des distributions nettement dissymétriques, une mesure corrective sera suggérée. Un exemple d'estimation de la concentration moyenne en plutonium du sol superficiel du Nevada Test Site, Nevada, U.S.A., sera consideré.

3.2 Simulation par RSS

Pour les simulations on a utilisé le langage R avec le package **RSSampling** crée par Bursa et al. 2018.

TAB. 3.1 – Moyenne et variance empirique de RSS et SRS et l'efficacité relative de RSS par rapport au SRS dans le cas du classement parfait avec distribution normale $\mathcal{N}(0,1)$.

n	$\bar{X}_{SRS}$	$\bar{X}_{RSS}$	σ^2_{SRS}	σ^2_{RSS}	éff$(\sigma^2_{SRS}, \sigma^2_{RSS})$
100	-0.03311767	0.1792694	0.9820492	0.9652943	1.017357297
1000	-0.01902246	0.09325392	1.020995	0.9153805	1.11537704
10000	0.001166947	0.2769816	0.9479426	0.665203	1.425042581

TAB. 3.2 – Efficacité relative de RSS par rapport au SRS pour un échantillon normal ,exponentielle ,gamma standard de taille 10000 avec k=3

Distribution	m=1	m=2	m=3
N$(0,1)$	1.180109	1.2	1.584722
exp(1)	1.206073	1.619588	1.226402
gamma	1.121302	1.505752	1.140202

TAB. 3.3 – Efficacité relative de RSS par rapport au SRS pour un échantillon normal ,exponentielle ,gamma standard de taille 1000

$K=3$

Distribution	m=1	m=2	m=3
N$(0,1)$	1.095139	1.470619	1.113598
exp(1)	1.215282	1.631953	1.235766
gamma	1.041989	1.399246	1.059552

3.3 Test d'ajustement

3.4 Applications sur les mesures de risque

3.4.1 Introduction

Au cours des siècles, de nombreux évènement extrêmes à travers le monde, tels les catastrophes naturelles (tremblements de terre, inondations, violents

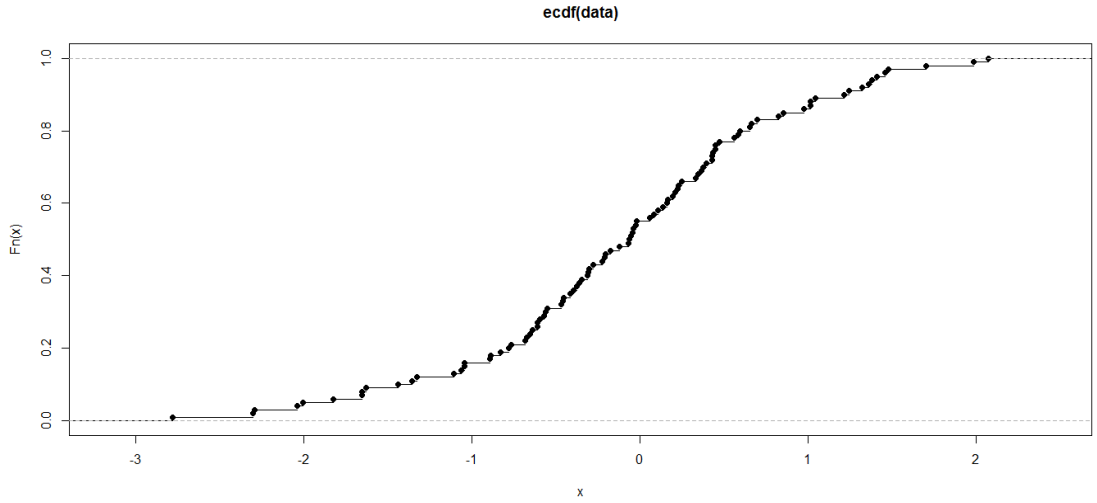

FIG. 3.1 – Fonction de distribution cumulative empirique de SRS pour n=100

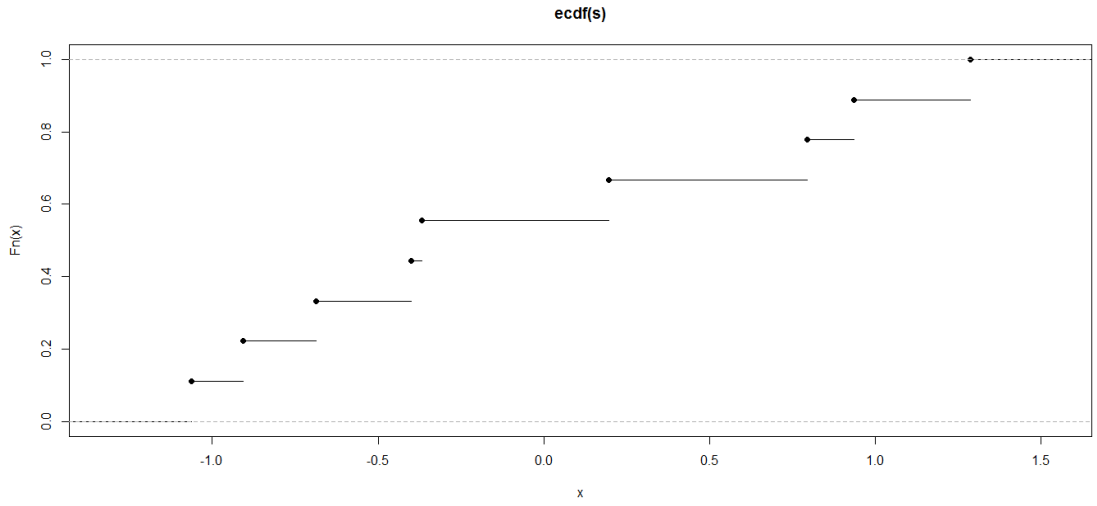

FIG. 3.2 – Fonction de distribution cumulative empirique de RSS pour n=100

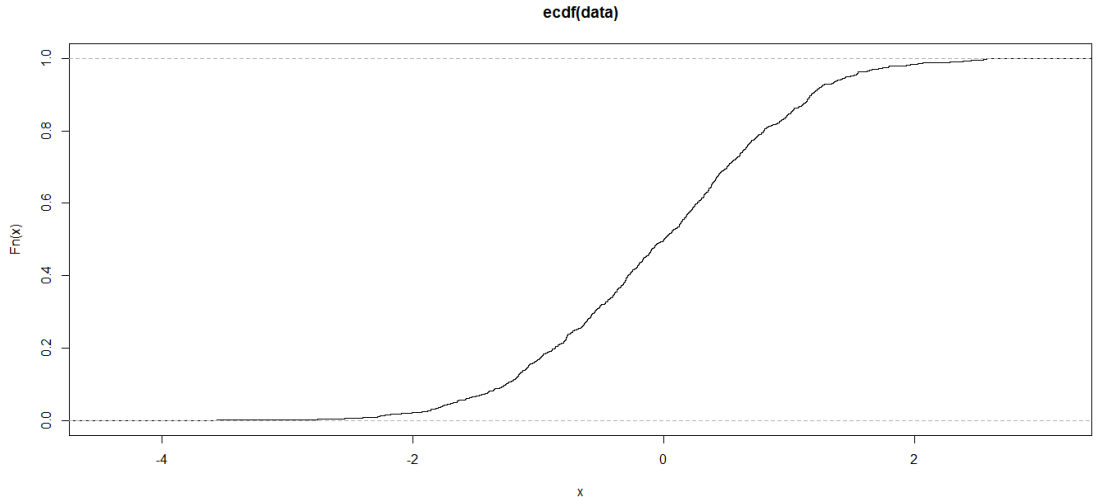

FIG. 3.3 – Fonction de distribution cumulative empirique de SRS pour n=1000

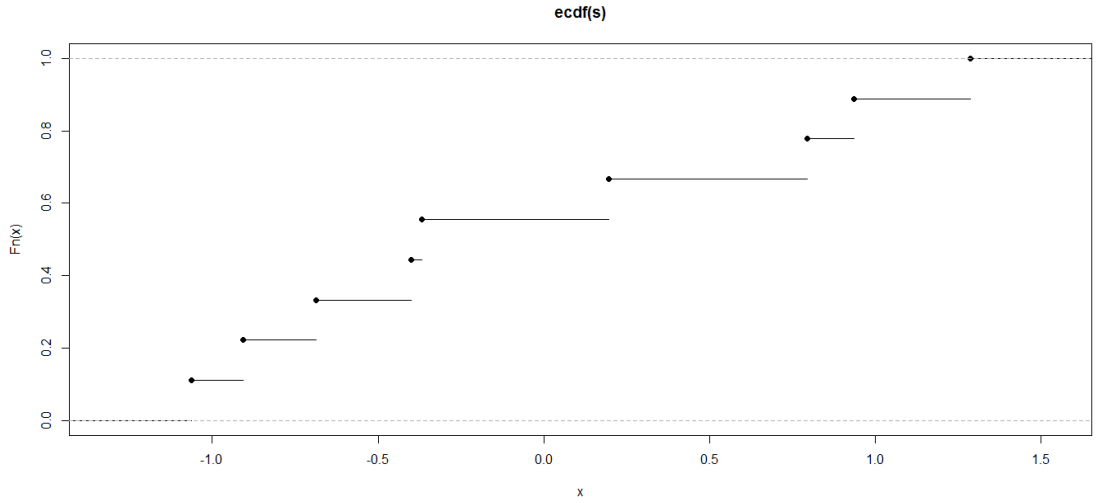

FIG. 3.4 – Fonction de distribution cumulative empirique de RSS pour n=1000

TAB. 3.4 – Calcul du quantile (q) à partir d'une probabilité $(1-\alpha) = 0.95$

Distibution	Quantile théorique	Quantile empirique de SRS	Quantile empirique de RSS
N(0,1)	1.644854	1.664556	1.630357
unif(0,1)	0.95	0.9024051	0.93010552
exp(1)	2.995732	2.959956	2.997107

TAB. 3.5 – Calcul du quantile (q) à partir d'une probabilité $(1-\alpha) = 0.975$

Distibution	Quantile théorique	Quantile empirique de RSS	Quantile empirique de RSS
N(0,1)	1.959964	1.839934	1.970215
unif(0,1)	0.975	0.9077724	0.97157059
exp(1)	3.688879	3.728473	3.669744

orages, . . .) ou les accidents liés à l'activité humaine (circulation routière, guerres, incendies, pollution industrielle ou nucléaire, . . .), ont représenté une menace réelle à l'homme, son économie et son environnement. Par conséquent, il est impératif de prendre en compte et d'anticiper les possibilités de survenance de tels phénomènes, afin d'en limiter les effets négatifs. Un risque est défini par la probabilité d'apparition d'un évènement rare et par l'ampleur de ses conséquences (probabilité et impact). Un plan d'atténuation du risque s'attachera donc à maitriser leur probabilité de survenance mais aussi à réduire leur impact. Historiquement, la notion de risque était liée à celle de probabilité. Apparue ou XVII siècle dans l'analyse des jeux de hasard, elle fut appliquée au XVIII-siècle par les assureurs maritimes, pour devenir ensuite une partie intégrante des schémas de prise de décision rationnelle associée à toute alternative de probabilités de succès ou d'échec. En 1921, Frank Knight (1921)[**?**] introduisit dans son livre « Risk, Uncertainty and Profit » la distinction entre risque (pouvant être calculé) et incertitude.

Ce chapitre a pour but de synthétiser les connaissances principales sur les mesures de risque utilisées pour le calcul des réserves de solvabilité, qui sont par excellence la Value-at-Risk (VaR) et la Tail Value-at-Risk (TVaR). Afin d'étudier ces mesures de risque à proprement parler, nous commençons par aborder le concept de mesure de risque cohérente, comonotone, et monétaire.L'essentiel des développements présentés dans ce chapitre proviennent du livre de Charpentier(2008)[6] et Ruszczynski (2006)[28].

TAB. 3.6 – Calcul du quantile (q) à partir d'une probabilité $(1-\alpha) = 0.99$

Distibution	Quantile théorique	Quantile empirique de RSS	la valeur de RSS
N(0,1)	2.326348	2.275079	2.30229
unif(0,1)	0.99	0.9343127	0.9902451
exp(1)	4.60517	4.56459	4.583178

TAB. 3.7 – Variance et précision relative de RSS par rapport au SRS dans le cas du classement parfait avec distribution normale $\mathcal{N}(2,3)$

n	σ^2_{SRS}	σ^2_{RSS}	éff$(\sigma^2_{SRS}, \sigma^2_{RSS})$
100	10.9238	9.830219	1.11124686
1000	9.005856	8.982007	1.0026552
10000	9.231003	6.896217	1.3385604

3.5 Rappel sur les risques

Le risque est un mot avec des implications diverses. Certaines personnes définissent le risque différemment des neutres. Ces désaccords cause de graves confusions dans le domaine de l'évaluation des risques et de leur gestion.

Définition 3.1 (le risque) *Un risque est une v.a. définie sur l'espace mesurable (Ω, F) désigné par X. Il représente la perte nette finale d'une position (éventualité) actuellement détenue. Lorsque $X > 0$, on dit qu'il y a une perte, la classe de ces v.a. sur (Ω, F) est noté par ϕ.*

Types de risque

Le risque est généralement classé en trois grandes catégories :

Risque de marché :

Il s'agit du risque de perte lié à l'évolution des niveaux ou à la volatilité des prix du marché. Les différents facteurs de risques liés au marché sont les taux, les cours de change, les cours des actions et les prix des matières premières. Toute variation de ces données a un impact sur les positions et les portefeuilles. Il s'agit du principal champ d'utilisation de la VaR.

Risque de liquidité :

- Il est composé du risque de liquidité d'actifs (asset liquidity risk) et du risque de liquidité de financement (cash flow risk).
- L'Asset Liquidity Risk est le risque de ne pouvoir vendre à son prix un titre financier.

TAB. 3.8 – Test d'ajustement

teste	100	100	1000	1000	10000	10000
	SRS	RSS	SRS	RSS	SRS	RSS
Student	0.1944	0.7855	0.7925	0.9207	0.9324	0.9363
Wilcoxon	0.2675	0.4258	0.8233	0.9102	0.9363	1

Il peut se traduire, soit par une impossibilité effective de le vendre, soit par l'abaissement de la valeur de marché d'une société par rapport à ses filiales.

- Le Cash Flow Risk est lié au fait que les banques reçoivent majoritairement des dépôts à court terme de leurs clients et font des prêts à moyen et long terme. Il peut donc se créer un décalage entre les sommes prêtées et les sommes disponibles, car ces dernières peuvent être insuffisantes. Dans ce cas on parle de manque de liquidités.

Risque de crédit :

Il résulte de l'incertitude quant à la possibilité ou la volonté des contreparties ou des clients de remplir leurs obligations. Il existe un risque pour une banque, dès qu'elle se met en situation d'attendre une entrée de fonds de la part d'un client ou d'une contrepartie du marché.

3.5.1 Définitions et propriétés

Définition 3.2 (Fonction P&L et fonction de perte) *Considérons P_t comme la valeur d'un portefeuille ou d'une position à la date t. Alors la variation de la valeur de ce portefeuille pour une période $[t, t+T]$, appelée fonction P&L (profit and loss), est :*

$$\Delta P_t = P_{t+T} - P_t.$$

et

$$X_i \simeq -\Delta P_t.$$

est appelé la fonction de perte.

Soit (Ω, F, P) un espace de probabilité et X_i soit une fonction mesurable à valeurs réelles définie sur Ω.

Soit $\mathcal{L}^\infty(\Omega, F, P)$ l'ensemble des variables aléatoires X_i définies sur Ω telles que

$$\|X_i\|_\infty = \sup_{w\in\Omega} |X_i| < \infty.$$

Soit $M \subset \mathcal{L}^\infty$ un cône convexe, c'est à dire que pour $X_1 \in M$ et $X_2 \in M$ on a $X_1 + X_2 \in M$ et $\lambda X \in M$ pour tout $\lambda > 0$.

M est l'ensemble de variables aléatoires réelles définies sur l'espace probabilisé (Ω,F, P). On interprète M comme un ensemble de pertes à horizon fixé.

Nous reprenons ici la définition d'une mesure de risque telle qu'elle est formalisée dans Charpentier A. Denuit M. (2004)[8]

Définition 3.3 *Mathématiquement, une mesure de risque ou un besoin en capital, d'une (v.a.) de perte X, est définie comme une fonction d'une perte aléatoire à un nombre réel.*

Définition 3.4 (mesure de risque) *On appelle mesure de risque une fonction R associant à X une valeur positive ou nulle telle que :*

$$\begin{array}{rcl} R & : & M \to \mathbb{R} \\ X & \to & R(X) \end{array} \tag{3.1}$$

Une mesure de risque $R(X)$ est interprétée comme le montant minimum qui, additionné à la perte X en début de période rend la couverture de X « acceptable ». C'est donc le capital dont doit disposer la compagnie pour faire face à une perte financière de montant X. Alors une mesure de risque, basée sur une v.a.de perte, est conçue pour indiquer à quel point la perte aléatoire est risquée. Dans le contexte actuariel, une mesure de risque est définie comme une fonction faisant correspondre à une perte aléatoire non négative une valeur réelle non négative.

On considère que la position de perte X est «dangereuse » si $R(X)$ est grand. On va interpréter $R(X)$ comme le cash à ajouter à une position de perte X pour qu'elle devienne acceptable pour le régulateur (i.e. pour que $R(X) > 0$).

3.5.2 Propriétés d'une mesure de risque

♢ Mesure de risque cohérente

Le concept de mesure de risque cohérente a été abordé pour la première fois par Artzner et al (1999.)[3]. Ce concept a été reformulé par la suite par différents auteurs. Nous utilisons la formulation proposée par Shapiro et Ruszczynski (2006)[**?**] pour des variables aléatoires mesurant des pertes. Ainsi, plus la variable aléatoire est grande, plus la mesure de risque est grande, plus le risque financier est important.

On dit que la mesure de risque R est :

Définition 3.5 –

– *La sous additivté :*

$$\forall X, Y \in M : R(X+Y) \leq R(X) + R(Y),$$

plus généralement

$$\forall X \in M \ R(nX) = R(X + ... + X) \leq nR(X), n = 1, 2, ...$$

– *L'homogénéité positive :*

$$\forall X \in M \ R(\lambda X) \leq \lambda R(X) \ \textit{pourtout} \ \lambda \in \mathbb{R}_+,$$

plus généralement

$$R(nX) \leq nR(X), \ n = 1, 2, ...$$

De même qu'une fusion ne crée pas risque supplémentaire $(R(\lambda X) \leq \lambda R(X))$*, une fusion sans diversification ne réduit pas le besoin global en capital.*

– *La monotonie :*

$$\forall X, Y \in M \quad X \geq Y \Rightarrow R(X) \geq R(Y).$$

Si les pertes encourues avec le risque X *sont toujours supérieures à celles obtenues avec* Y*, le besoin en capital pour* X *doit être supérieur à celui pour* Y*.*

– *L'invariance par translation :*

$$\forall X \in M \quad R(X+c) \leq R(X) + c \ \forall c \in \mathbb{R}. \tag{3.4}$$

Speciallement, nous avons $R(X + (X)) = R(X) - R(X) = 0$ *c'est-à-dire, en ajoutant* (X) *à la position initiale* X*, nous obtenons une position "neutre". On notera en particulier que* $R(X) = -\alpha$ *avec la convention que* $R(0) = 0$*.*

– *Convexe :*
La convexité implique que la diversification n'augmente pas le risque, car la valeur de risque du portefeuille diversifié $\lambda X + (1-\lambda)Y$ *est inférieure ou égale à la moyenne re-pondérer des différentes valeurs de risque.*
R *est dite convexe si pour tout* $\lambda \in [0, 1]$

$$\forall X, Y \in M \quad R(\lambda X + (1-\lambda) Y) \leq \lambda R(X) + (1-\lambda) R(Y). \tag{3.5}$$

Une mesure de risque convexe et positivement homogène est dite cohérente.

Mesure de risque monétaire

Le concept de mesure de risque monétaire (monetary mesure of risk) donne une interprétation concrète aux mesures de risque. Nous allons tout d'abord définir cette notion de mesure de risque monétaire, elle admet une interprétation simple : c'est la quantité d'actif sans risque qui, ajouté à une position, permet de rendre celle-ci acceptable par le régulateur.

Définition 3.6 *Soit une mesure de risque R de :$M \to \mathbb{R}$,*

- *R est dite monétaire si elle est monotone et invariante par translation,*
- *R est dite convexe si elle est monétaire,*
- *R est dite cohérente si elle est monétaire, homogène et sous-additive.*

Corollaire 3.1 *Si R est une mesure de risque monétaire, homogène et normalisée à 0 (i.e $R(0) = 0$), alors la convexité et la sous-additivité sont des notions équivalentes.*

Définition 3.7 (Région de risques acceptables) *R est une mesure de risque, on définit la région de risque acceptable pour la mesure comme*

$$A = \{X \in M, R(X) \leq 0\}. \tag{3.6}$$

Réciproquement, si A est une région de risque acceptable, la mesure de risque induite est

$$R(X) = \inf\{m, X - m \in A\} \tag{3.7}$$

Proposition 3.1 *R est une mesure de risque monétaire alors :*

- *est convexe si et seulement si A est convexe,*
- *est positivement homogène si et seulement si A est un cône.*

Preuve: Pour le premier point, si R est convexe, alors A est convexe. Réciproquement, si A est convexe, soient X_1, X_2, m_1 et m_2 tels que $X_i - m_i \in$, pour $i = 1, 2$. Par hypothèse, pour tout λ , $\lambda[X_1 - m_1] + (1-\lambda)[X_2 - m_2] \in A$, c'est-à-dire que $R([X_1 - m_1] + (1 - \lambda)[X_2 - m_2] \leq 0$, soit, par la propriété d'invariance par translation, et par monotonie,

$$R(X_1 + (1 - \lambda)X_2) \leq \lambda m_1 + (1 - \lambda)m_2 \tag{3.8}$$

et ceci pour tout m_1 et m_2. Il suffit de le faire pour $m_i = (X_i)$. Pour le second point, si est un cône, alors pour tout $X - m \in A$, $\lambda(X - m) \in A$, pour tout $\lambda > 0$. Donc $R(\lambda X - \lambda m) \leq 0$, d'où $R(\lambda X) \leq \lambda R m$. Si $m = R(X)$, on en déduit que $R(\lambda X) \leq \lambda R(X)$. Et si $X - m \notin A$; alors $\lambda(X - m) \notin A$, et $R(\lambda X) > \lambda m$. On fait alors tendre m vers $R(X)$ pour avoir le résultat souhaité. On dispose du théorème de représentation suivant ■

Théorème 3.1 *R est une mesure de risque monétaire convexe si et seulement si pour tout X bornée ($X \in L^{\infty}$)*

$$R(X) = \max_{Q \in M} \{E_Q - \alpha(Q)\}, \tag{3.9}$$

où est l'ensemble des mesures additives et normalisées à 1, et

$$\alpha(Q) = \sup_{X \in A} \{E_Q(X)\} \tag{3.10}$$

où A est l'ensemble d'acceptation associé à R. On notera que contient plus que des mesures de probabilité.

Proposition 3.2 *Si R est une mesure de risque convexe, les trois propriétés suivantessont équivalentes*

1. *est fortement cohérente,*
2. *est additive pour des risques comonotones,*
3. *est une mesure de corrélation maximale.*

Proposition 3.3 *Une mesure de risque cohèrente R est additive pour des risques comonotones si et seulement s'il existe une fonction décroissante positive g sur $[0;1]$ telle que*

$$R(X) = \int_0^1 g(t) F_X^{-1}(1-t)\, dt \tag{3.11}$$

où $F_X(x) = P(X \leq x)$.

Il existe de nombreuses façons de mesurer le risque, la mesure la plus répandue est la Value-at-Risk (VaR) établie par JP Morgan (1994).

3.5.3 Value-at-Risque (VaR)

La VaR est utilisé aussi bien par les institutions financières et les régulateurs, que par les entreprises non financières. Les institutions financières ont été les premières à utiliser cet outil. En effet, la diversification des risques financiers, la complexité des nouveaux instruments financiers et l'évolution de la régulation ont poussé les institutions financières à mettre en place des systèmes centralisés de surveillance et de management des risques. De leur côté, les réglementations doivent évaluer ces risques financiers afin d'imposer aux institutions financières un niveau minimal de capitaux disponibles.

Par ailleurs, le management centralisé des risques est utile à toute entreprise exposée à des risques financiers. Les entreprises non financières, comme

les multinationales par exemple, utilisent quant à elles la VaR pour se prémunir contre le risque de change.

La Value-at-Risk permet de mesurer différents risques, sur différents marchés (marché des changes, marché financier, marché des produits dérivés), et pour différents actifs à risque (change, actions, obligations, options, etc.).

L'objectif de la VaR est de fournir une mesure du risque total de portefeuille. Par conséquent, la VaR doit tenir compte des effets de levier et de diversification. En effet, la diversification d'un portefeuille de titres ou d'actifs permet en variant les types de placements, soit de réduire le risque pour un niveau de rentabilité donné, soit d'améliorer la rentabilité pour un niveau de risque donné. Pour un groupe la diversification permet de réduire le risque de volatilité des résultats. La VaR mesure donc différents risques financiers.

Présentation générale

« Combien, au maximum, je peux perdre sur cet investissement ? » est la question que chaque investisseur s'est probablement posé en investissant dans un ou plusieurs actifs risqués. Pour y répondre, le concept de la **VaR** semble être une bonne alternative. Simple et utilisée par tous, elle offre également l'avantage d'être une mesure prospective du risque. La définition générale utilisée par les praticiens est la suivante : « **La Value-at-Risk** correspond au montant des pertes qui ne devrait pas être dépassé pour un niveau de c^{2}onfiance donné et sur un horizon temporel fixé. »

3.5.4 Définition de la Value-at-Risk

Quantile d'ordre α

Définition 3.8 *Soient X est une v.a.r et F sa fonction de répartition. On appelle quantile ou fractile d'ordre α, le nombre x défini par :*

$$x_\alpha = \inf \{x \in \mathbb{R} \mid F(x) \geqslant \alpha\}, avec\ \alpha \in [0,1].$$

Remarque 3.1 *Si F est strictement croissante et continue, alors x_α est l'unique nombre réel tel que*

$$F(x_\alpha) = \alpha$$

Statistiquement, la VaR se définie pour un taux de couverture de niveau $\alpha \in [0,1]$ comme étant le quantile de niveau, de la distribution de profits et de pertes (P&L) relatif à un portefeuille d'actifs pour une période donnée. Plusieurs formulations ont été proposées.

Nous reprenons la formulation la plus usuelle proposée par Charpentier (2008)[6] portant sur des variables aléatoires mesurant des pertes.

Définition 3.9 (VaR) *La Value-at-Risk correspondante au portefeuille d'actifs, notée* $VaR(X;\alpha)$*, est le quantile d'ordre* α *du montant du portefeuille* X*.*

$$VaR(X;\alpha) = x_\alpha. \tag{3.12}$$

Où,

$$F(x) = P(X \leqslant x_\alpha) = \alpha.$$

D'où :

$$VaR(X;\alpha) = \inf\{x, P(X \leqslant x_\alpha) \geqslant \alpha\} = F_X^{-1}(\alpha) = Q(\alpha). \tag{3.13}$$

Pour un calcul de risques financiers, on est intéressé par de grandes valeurs de Typiques :
$\alpha = 0,05$ ou $\alpha = 0,01$.

3.5.5 Représentation graphique de la VaR

Nous avons utilisé les cours quotidiens de la compagnie d'assurance d'AXA (présenté dans l'annexe B) qui sont observés sur la période du 25/10/2007 au 23/10/2009.

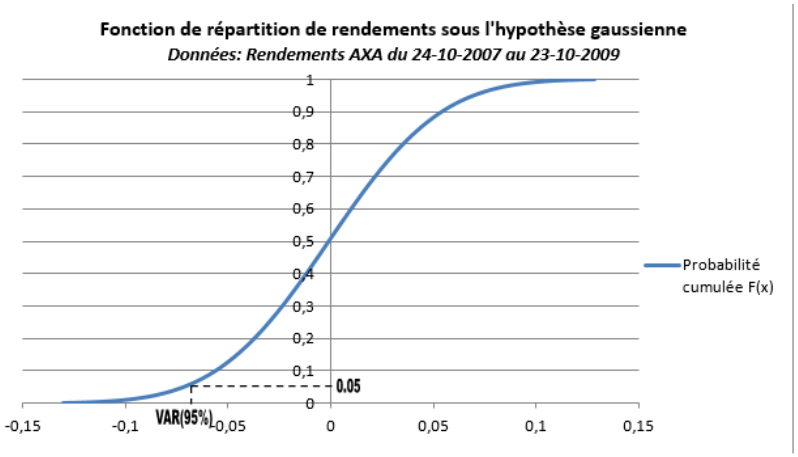

FIG. 3.5 – Fonction de répartition de rendements sous l'hypothèse gaussienne

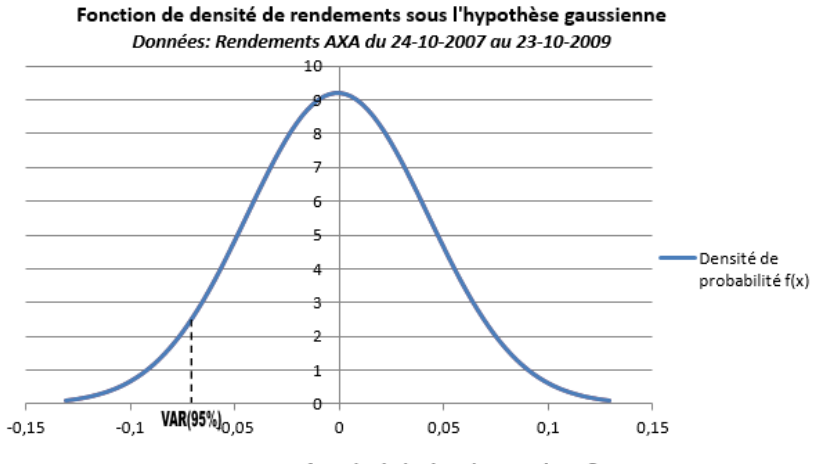

FIG. 3.6 – La V aR, un fractile de la distribution des P&L.

Remarque 3.2 *On notera que* $VaR[X;\alpha]$ *est une fonction croissante en* α*, contrairement à certains auteurs qui notent* $VaR[X;\alpha]$ *le quantile d'ordre* $1-\alpha$*.*

Commençons par noter que la VaR *est stable par transformation croissante (non-linéaire) : quel que soit le niveau de probabilité* $\alpha \in (0,1)$ *et la fonction croissante et continue* g*,*

$$VaR[g(X);\alpha] = g\left(VaR[X;\alpha].\right)$$

La VaR d'un portefeuille dépend essentiellement de 3 paramètres :

* ***La distribution des profits et des pertes (P&L) du portefeuille ou de l'actif :***

  Souvent cette distribution est supposée gaussienne, mais beaucoup d'acteurs financiers utilisent des distributions historiques. La difficulté réside alors dans la taille de l'échantillon historique : s'il est trop petit, les probabilités de pertes élevées sont peu précises, et s'il est trop grand, la cohérence temporelle des résultats est perdue, car on

  compare des résultats non comparables. Les données de (P&L) à partir desquelles on calcule une VaR sont généralement exprimées sous forme de rendements.

* ***Le niveau de confiance choisi (95% ou 99 % en général) :***

C'est la probabilité que les pertes éventuelles du portefeuille ou de l'actif ne dépassent pas la VaR.

* ***L'horizon temporel choisi (période de détention de l'actif) :***

Ce paramètre est très important car plus l'horizon est long plus les pertes peuvent être importantes.

3.5.6 Au delà de la VaR

Pour remédier à ces défauts de la VaR, d'autre mesures ont été proposés. Nous en présentons trois très proches, dont l'idée commune est de quantifier le risque lorsque la VaR est dépassé.

- La Tail Value-at-Risk ou TVaR est la moyenne des VaR de niveau supérieur à α ;
- La Conditional Tail Expectation ou CTE représente la perte attendue sachant que la VaR est dépassée ;
- L'Expected Shortfall, ou ES au niveau α, c'est la prime stop-loss dont la rétention (ou franchise, ou la priorité) est fixé à VaR_α.

Nous définissons précisément ces mesures ci-dessous.

3.5.7 La Tail Value-at-Risk

Beaucoup d'auteurs et d'articles dans la littérature définissent la Tail Value-at-Risk (TVaR). C'est cependant l'approche de Charpentier (2008)[6] que nous retiendrons

Définition 3.10 (TVaR) *La Tail Value-at-Risk au niveau α, notée $TVaR(X;\alpha)$ est définie par :*

$$TVaR\left(X;\alpha\right) = \frac{1}{1-\alpha}\int_{\alpha}^{1} VaR\left(X;t\right)dt. \tag{3.14}$$

$\Longrightarrow$la Tail-VaR est la moyenne des VaR de niveau supérieur à α.
Notons que la TVaR est plus grande que la VaR correspondante.

Remarque 3.3 *Il existe une fonction de répartition $\widetilde{F_X}$ (transformée de Hardy-Littlewood de F_X), telle que pour tout α :*

$$\widetilde{F}_X^{-1}\left(\alpha\right) = TVaR\left(X;\alpha\right). \tag{3.15}$$

Si on note $\widetilde{X}$ une variable aléatoire de fonction de répartition F_X, on a :

$$TVaR\left(X;\alpha\right)=VaR\left(\widetilde{X};\alpha\right). \tag{3.16}$$

La $TVaR$ est donc la VaR de la transformée de Hardy-Littlewood de X. Notons que :

$TVaR[X;0]=E[X]$. Et comme

$$TVaR\left[X;\alpha\right]=\frac{1}{1-\alpha}\left\{\left[X\right]-\int_0^{\alpha}VaR\left[X;\xi\right]d\xi\right\}. \tag{3.17}$$

on en déduit que la Tail Value-at-Risk est une fonction croissante du niveau α.

En effet :

$$\frac{d}{d\alpha}TVaR\left(X;\alpha\right)=\frac{TVaR\left(X;\alpha\right)}{1-\alpha}-\frac{VaR\left(X;\alpha\right)}{1-\alpha}. \tag{3.18}$$

Et comme $\alpha\longmapsto VaR[X;\alpha]$ est une fonction croissante :

$$TVaR\left[X;\alpha\right]=\frac{1}{1-\alpha}\int_{\alpha}^{1}\underbrace{VaR\left[X;t\right]}_{\geq VaR[X;\alpha]}dt\geq VaR\left[X;\alpha\right], \tag{3.19}$$

on en déduit que $\frac{d}{d\alpha}TVaR[X;\alpha]\geq 0$, et :

$$TVaR\left[X;\alpha\right]\geq TVaR\left[X;0\right]=E\left[X\right]. \tag{3.20}$$

Proposition 3.4 *La TVaR est cohérente.*

3.6 Application sur l'indice boursier CAC40

3.6.1 Introduction

Le CAC 40 est un indice boursier regroupant les 40 plus importantes capitalisations boursières françaises cotées à la bourse de Paris. « CAC » signifie « Cotation Assistée en Continu ».

Le CAC 40 est utilisé comme un indicateur de l'évolution économique des grandes entreprises françaises. La liste des 40 entreprises qui y figurent est régulièrement mise à jour en vue de maintenir cette représentativité. L'indice est calculé en continu tous les jours et fait l'objet d'une mise à jour toutes les 30 secondes. Le CAC 40 fait parti du groupe de cotation Euronext.

TAB. 3.9 – La Value at Risque VaR et l'Expected shortfall ES au niveau de probabilité α

α	0.9	0.95	0.975	0.99
VaR_{SRS}	4113.814	4381.996	4607.378	4854.379
VaR_{RSS}	4244.293	3703.275	4872.88	4534.384
ES_{SRS}	4525.179	4734.422	4932.403	5094.315
ES_{RSS}	4260.74	4310.69	4260.74	4606.8

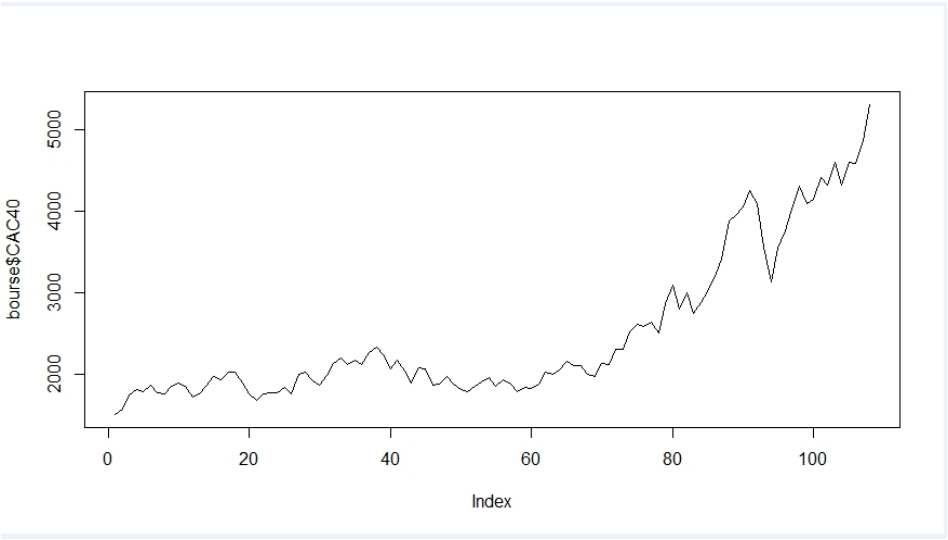

FIG. 3.7 – L'indice CAC 40 de 1991 jusqu'a 1999

Chapitre 4

Conclusion générale

Dans ce travail, nous avons donnée une étude sur la méthode d'échantillonnage classé depuis la naissance de l'idée original .

La méthode est d'abord présentée dans le cas d'une population univariée dont on cherche à estimer l'espérance et la variance. On montre que sous des hypothèses assez faibles, la moyenne expérimentale d'un échantillon classé est sans biais et que sa variance est inférieure à celle de l'estimateur traditionnel basé sur un échantillonnage aléatoire simple de même taille. Une généralisation de la méthode d'échantillonnage classé est ensuite considérée dans le cas d'une population à deux caractères aléatoires. On montre que dans les mêmes conditions que pour le cas univarié, l'efficacité de l'estimateur fondé sur l'échantillon classé bivarié est supérieure à celle de la moyenne bivariée calculée à partir d'un échantillonnage aléatoire simple. Les inférences statistiques basées sur RSS sont étudiées à partir de l'estimation prévue de la moyenne empirique à de nombreuses procédures plus compliquées telles que les inférences sur les moyennes, les variances, les quantiles et fonctions de densité, les tests sans distribution.

Nous avons montré la performance et l'efficacité relative de cette application dans ce cas l'estimateur RSS est plus efficace que l'estimateur de SRS par des moyens de simulations des variables aléatoires usuelles et avec une application sur des données réelles, surtout pour la valeur du quantile empirique.

Bibliographie

[1] Al-Saleh, M. F. & Zheng, G. (2002). Estimation of bivariate characteristics using ranked set sampling. Australian and New Zealand Journal of Statistics, 44, 221-232.

[2] Arnold, B. C., Balakrishnan, N., and Nagaraja, H. N. (1992). A First Course in Order Statistics. New York : Wiley.

[3] Artzner P. Delbaen F. Eber J -M and Heath D. (1999). Coherent measures of risk, Mathematical Finance, 9 :203-228.

[4] Bohn, L. L. & Wolfe, D. A. (1992). Nonparametric two-sample procedures forranked-set samples data. Journal of the American Statistical Association, 87, 552-561.

[5] Charpentier A. Denuit M. (2004). Mathématiques de l'assurance nonvie. Tome 1 : principes fondamentaux de théorie du risqu

[6] Charpentier A. (2008). Value at risk et probabilitée de ruine, entre vaccination et banque d'affaires. Risques, 76 :103-106.

[7] Z. Chen. (2000). On ranked-set sample quantiles and their applications. Journal of Statistical Planning and Inference, 83 :125–135.

E.Parzen (1979) Nonparametric statistical data modelling .J.Am.Stat.Assoc 74, page 105-131

[8] Charpentier A. Denuit M. (2004). Mathématiques de l'assurance nonvie. Tome 1 : principes fondamentaux de théorie du risque. Economica.

[9] David, H. A. (1981). Order Statistics, deuxième édition. Wiley, New York.

[10] Dell, T. R. and Clutter, J. L. (1972). Ranked set sampling theory with order statistics background. Biometrics, 28, 545-555.

[11] Hedayat, A. S. and Sinha, B. K. (1992). Design and Inference in Finite Population Sampling. New York : Wiley.

[12] Knight F H. 1921. Risk, incertainty, and Profit. Houghton Mifflin Company, Boston.

[13] McIntyre, G. A. (1952). A method for unbiased selective sampling, using ranked sets. Australian Journal of Agricultural Research, 3, 385-390.

[14] Patil G. P., Sinha A. K. & Taillie C. (1999). Ranked set sampling : A bibliography. Environmental and Ecological Statistics, 6, 91-98.

[15] E.Parzen (1979) Nonparametric statistical data modelling .J.Am.Stat.Assoc 74, page 105-131

means of ordering. Annals of the Institute of Statistical Mathematics, 20, 1-31.

[16] Stokes, S. L. (1977). Ranked set sampling with concomitant variables. Communications in Statistics—Theory and Methods, 6, 1207-1211.

[17] Stokes, S. L. & Sager, T. W. (1988). Characterization of a ranked-set sample with application to estimating distribution functions. Journal of the American Statistical Association, 83, 374-381.

[18] R. J. Serfling.(1980). Approximation theorems of mathematical statistics. John Wiley & Sons, New York.

[19] B. W. Silverman. (1986). Density estimation for statistics and data analysis. Chapman and Hall, London, New York.

[20] Stokes, S. L. (1977). Ranked set sampling with concomitant variables. Communication in Statistics : Theory and Methods 6, 1207-1211.

[21] Sukhatme, P. V. and Sukhatme, B. V. (1970). Sampling Theory of Surveys with Applications. Ames : Iowa State University Press

[22] Stokes, S. L. (1977). Ranked set sampling with concomitant variables. Communication in Statistics : Theory and Methods 6, 1207-1211.

[23] Sukhatme, P. V. and Sukhatme, B. V. (1970). Sampling Theory of Surveys with Applications. Ames : Iowa State University Press

[24] Takahasi, K. & Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. Annals of the Institute of Statistical Mathematics, 20, 1-31.

[25] Tikkiwal, B. D. (1960). On the theory of classical regression and double sampling estimation. Journal of the Royal Statistical Society, Series B 22, 131-138.

[26] Ruszczynski A and Shapiro A. (2006) Conditional risk mappings. Mathematics of Operations Research, 31(3) :544-561, August .

[27] M.Rosenblatt (1956). Remarks on some nonparametric estimates of a density function .Ann Math. Stat.27,page 832-837

[28] Ruszczynski A and Shapiro A. (2006) Conditional risk mappings. Mathematics of Operations Research, 31(3) :544-561, August .

[29] Wolfe, D. A. (2004). Ranked set sampling : An approach to more effecient data collection. Statistical Science, 19, 636-643