



**HAL**  
open science

# Développements d'Edgeworth de deux estimateurs d'une proportion de mesures—Edgeworth expansions for two point estimators of the proportion of measures

Frédéric Bertrand, Myriam Maumy

► **To cite this version:**

Frédéric Bertrand, Myriam Maumy. Développements d'Edgeworth de deux estimateurs d'une proportion de mesures—Edgeworth expansions for two point estimators of the proportion of measures. *Comptes Rendus. Mathématique*, 2007, 345 (7), pp.399-404. 10.1016/j.crma.2007.07.017. hal-00184244

**HAL Id: hal-00184244**

**<https://hal.science/hal-00184244>**

Submitted on 31 Oct 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Développements d'Edgeworth de deux estimateurs d'une proportion de mesures

Frédéric Bertrand<sup>1</sup>, Myriam Maumy<sup>1</sup>  
fbertran@math.u-strasbg.fr, mmaumy@math.u-strasbg.fr

<sup>1</sup> Institut de Recherche Mathématique Avancée, Université Louis Pasteur,  
7, rue René Descartes, 67084 Strasbourg Cedex, France

## Résumé

Dans ce travail, nous établissons des développements d'Edgeworth de deux estimateurs d'une proportion de mesures  $\pi$ . Le premier est construit à partir de la méthode du maximum de vraisemblance et le second à l'aide des estimateurs sans biais et de variance minimale. Ces résultats sont motivés par leurs applications dans les laboratoires de contrôle ou dans l'industrie chimique, pharmaceutique, agroalimentaire, etc.

## Abstract

### Edgeworth expansions for two point estimators of the proportion of measures

In this paper, we establish Edgeworth expansions for two point estimators of a proportion of experimental results  $\pi$ . The first one is based on the maximum likelihood method and the second one is the minimum variance unbiased estimator. These results are motivated by their counterparts in the official quality laboratories and in the industries of various sectors, namely: chemistry, pharmacy, food processing, etc.

## Abridged English Version

### Issues and aims

One burning issue that commonly arises in official quality laboratories or in the industries of various sectors, namely chemistry, pharmacy, food processing, etc, is the estimation of the proportion of experimental results that lie within two limits.

The litterature is rather well furnished on this subject : let's cite among all these works [2] and [9] for further details. Assume that a sequence of experimental results that follow a normal distribution with unknown mean  $\mu$  and unknown variance  $\sigma^2 > 0$  and that the proportion of experimental results that lie within the limits of the interval  $[L, U]$ , where  $L$  stands for Lower and  $U$  for Upper, denoted by  $\pi$ , is the parameter of interest. For instance, in official quality laboratories,  $[L, U]$  is the the range of the acceptable values and then  $\pi$  is the proportion of valid units.

In this Note, we introduce two possible point estimators for the proportion  $\pi$  and their Edgeworth expansions which will allow us to define confidence intervals that will be studied in another work.

The small sample sizes of experimental results, often ranging between 5 and 20, and therefore the fact that usual asymptotic normal approximations are irrelevant accounts for the use of these Edgeworth expansions.

### Notations and preliminary results

Let  $\{X_i : 2 \leq i \leq n\}$  be a sequence of experimental results that follow a normal distribution with mean  $\mu$  and variance  $\sigma^2 > 0$ . We denote by  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$  the sample mean and by  $s_c^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

the sample variance. We introduce the proportion  $\pi$  defined by :

$$\pi := \pi(\mu, \sigma^2) = \mathbb{P}[L \leq X \leq U] = \Phi\left(\frac{U - \mu}{\sigma}\right) - \Phi\left(\frac{L - \mu}{\sigma}\right), \quad (0.1)$$

where  $\Phi$  stands for the cumulative distribution function of the standard normal law,  $U$  and  $L$ , with  $L < U$ , are the “limits of acceptance”, whose values depend on the active regulatory norm.

Using (0.1) one can put forward two point estimators of  $\pi$  : the first, denoted by  $\hat{\pi}$  and defined by the Equation (2.2), is based on the maximum likelihood theory (MV) and the second, denoted by  $\tilde{\pi}$  and defined by the Equation (2.3), is unbiased with minimal variance (SBVM). Theorems 2.1 and 2.2 state that these two point estimators renormalized by  $\sqrt{n}$  are asymptotically normal; Theorem 2.2 is one of the results of this Note and its proof can be found in [1].

## Small sample simulations

We now provide two displays (Fig. 1) : the first plots  $\sqrt{n}(\hat{\pi} - \pi)$  and the second plots  $\sqrt{n}(\tilde{\pi} - \pi)$ , where  $\pi = 0,95$ ,  $L = -0,15$ ,  $U = 0,15$ ,  $\mu = 0$  and, following the decreasing order of the graphs, the density function for  $n = 10; 15; 20$  and for the normal limit law as  $n$  tends to infinity. These estimates of the density functions were computed with the Maple 10.0 software [8] using an Epanechnikov kernel and 10 000 simulated values of each of the two estimators.

We also simulated  $r = 10\,000$  values of  $\hat{\pi}$  and  $\tilde{\pi}$  for  $\pi = 0,95$ ,  $L = -0,15$ ,  $U = 0,15$ ,  $\mu = 0; 0,10$ ,  $n = 5; 10; 15; 20$ . The Table 1 sums up the results computed this way, using the fact that, if  $(x_1, \dots, x_r)$  is a  $r$ -sample, then Bias means Bias, Variance =  $\frac{1}{r} \cdot \sum_{i=1}^r (x_i - \bar{x})^2$ , MSE = Variance + Bias<sup>2</sup>, Asymétrie =  $\frac{1}{r} \cdot \sum_{i=1}^r (x_i - \bar{x})^3 / \text{Variance}^{3/2}$ , Asymétrie means Skewness, and Aplatissement =  $\frac{1}{r} \cdot \sum_{i=1}^r (x_i - \bar{x})^4 / \text{Variance}^2$ , Aplatissement means Kurtosis. The values of the skewness and the kurtosis lie far apart from the values of a normal distribution and therefore accounts for the use of the Edgeworth expansions whose computations is the aim of Section 4.

## Edgeworth and Cornish-Fisher expansions

Eventhough computing the Edgeworth expansions of  $\hat{\pi}$  and  $\tilde{\pi}$  is a fairly long task, even trickier for  $\tilde{\pi}$  than for  $\hat{\pi}$ , this work was completed and is detailed in the preprint [1].

The Theorems 4.1 and 4.2 state, with the Equations (4.3) and (4.4), the Edgeworth expansions of the two estimators  $\hat{\pi}$  and  $\tilde{\pi}$ .

These results lead to Cornish-Fisher expansions of the quantiles of the two estimators  $\hat{\pi}$  and  $\tilde{\pi}$  given by the Equation (5.1).

Yet the true values of the parameters  $\mu$  and  $\sigma$  of the normal distribution, whose values are required to compute the Cornish-Fisher expansions, are unknown. Bootstrap techniques can be used to solve that issue. This will be studied by the authors in an upcoming paper.

## 1. Motivation

Un des problèmes des laboratoires de contrôle ou de l'industrie chimique, pharmaceutique, agroalimentaire, etc, est l'estimation de la proportion d'une population de mesures se situant entre deux limites. La littérature est relativement abondante sur ce sujet : citons par exemple [2] et [9] pour de plus amples détails. Supposons qu'une série de mesures suive une loi normale de moyenne  $\mu$  et de variance  $\sigma^2 > 0$  inconnues et que la proportion des mesures appartenant à l'intervalle  $[L, U]$ ,  $L$  pour Lower et  $U$  pour Upper, notée  $\pi$ , est le paramètre d'intérêt. Par exemple, dans les laboratoires de contrôle,  $[L, U]$  représente l'étendue des valeurs produites acceptables et  $\pi$  est alors la proportion d'unités conformes. Dans cette Note, nous proposons deux estimateurs possibles pour la proportion  $\pi$  et établissons des développements

d'Edgeworth qui permettent de construire des intervalles de confiance qui seront étudiés par ailleurs. Le recours à ces développements se justifie par la faible taille des séries de mesures utilisées, souvent comprise entre 5 et 20, et par conséquent, par le fait que les approximations asymptotiques ne sont plus valides dans ce cas. La démonstration des Théorèmes 2.2, 4.1 et 4.2 est détaillée dans la prépublication [1].

## 2. Notations et résultats préliminaires

Soit  $\{X_i : 2 \leq i \leq n\}$  une série de mesures qui suit une loi normale de moyenne  $\mu$  et de variance  $\sigma^2 > 0$ . Désignons par  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$  la moyenne empirique et  $s_c^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$  la variance empirique corrigée. Introduisons la proportion  $\pi$  qui se définit par :

$$\pi := \pi(\mu, \sigma^2) = \mathbb{P}[L \leq X \leq U] = \Phi\left(\frac{U - \mu}{\sigma}\right) - \Phi\left(\frac{L - \mu}{\sigma}\right), \quad (2.1)$$

où  $\Phi$  désigne la fonction de répartition d'une loi normale centrée réduite,  $U$  et  $L$ , avec  $L < U$ , sont les « limites d'acceptation », dont les valeurs dépendent de la norme utilisée.

À partir de (2.1), deux estimateurs de  $\pi$  sont proposés : le premier, noté  $\hat{\pi}$ , est basé sur la méthode du maximum de vraisemblance (MV) et le second, noté  $\tilde{\pi}$ , est sans biais et de variance minimale (SBVM).

La propriété de l'invariance du maximum de vraisemblance permet de définir l'estimateur  $\hat{\pi}$  ainsi :

$$\hat{\pi} := \hat{\pi}\left(\bar{X}, \frac{s_c^2}{k^2}\right) = \Phi\left(k \frac{U - \bar{X}}{s_c}\right) - \Phi\left(k \frac{L - \bar{X}}{s_c}\right) = g_{MV}\left(\bar{X}, \frac{s_c^2}{k^2}\right) \quad \text{où} \quad k^2 = \frac{n}{n-1}. \quad (2.2)$$

**Théorème 2.1.**  $\hat{\pi} - \pi$  renormalisé par  $\sqrt{n}$  est asymptotiquement gaussien de variance  $\sigma_{MV}^2$  (cf [6]).

L'expression de l'estimateur  $\tilde{\pi}$  (cf [7]) s'écrit :

$$\begin{aligned} \tilde{\pi} := \tilde{\pi}\left(\bar{X}, \frac{s_c^2}{k^2}\right) &= (\tilde{\pi}_1 - \tilde{\pi}_2)\left(\bar{X}, \frac{s_c^2}{k^2}\right) & \text{où} \quad \tilde{\pi}_i\left(\bar{X}, \frac{s_c^2}{k^2}\right) &= \begin{cases} 0 & \text{si } b_i \leq 0 \\ 1 & \text{si } b_i \geq 1 \\ B_{b_i}\left(\frac{n-2}{2}, \frac{n-2}{2}\right) & \text{sinon,} \end{cases} \\ &= g_{SBVM}\left(\bar{X}, \frac{s_c^2}{k^2}\right) & & \end{aligned} \quad (2.3)$$

où  $B_{b_i}\left(\frac{n-2}{2}, \frac{n-2}{2}\right) = \frac{\Gamma(n-2)}{\Gamma(\frac{n-2}{2})\Gamma(\frac{n-2}{2})} \int_0^{b_i} u^{\frac{n-4}{2}} (1-u)^{\frac{n-4}{2}} du$ ,  $b_1 = \frac{1}{2} \left(1 + k \frac{\bar{X} - L}{s_c \sqrt{n-1}}\right)$  et  $b_2 = \frac{1}{2} \left(1 + k \frac{\bar{X} - U}{s_c \sqrt{n-1}}\right)$ .

**Théorème 2.2.**  $\tilde{\pi} - \pi$  renormalisé par  $\sqrt{n}$  est asymptotiquement gaussien de variance  $\sigma_{SBVM}^2 = \sigma_{MV}^2$ .

## 3. Quelques simulations pour $\hat{\pi}$ et $\tilde{\pi}$

Nous présentons ici deux graphiques (Fig. 1) : le premier pour  $\sqrt{n}(\hat{\pi} - \pi)$  et le second pour  $\sqrt{n}(\tilde{\pi} - \pi)$ , où  $\pi = 0,95$ ,  $L = -0,15$ ,  $U = 0,15$ ,  $\mu = 0$  et par ordre décroissant des maxima des courbes, la densité pour  $n = 10$ ,  $n = 15$ ,  $n = 20$  et la loi normale limite lorsque  $n$  tend vers l'infini. Ces estimations de la densité ont été obtenues à l'aide du logiciel Maple 10.0 [8] et au moyen d'un noyau d'Epanechnikov et de 10 000 réalisations de chacun des estimateurs.

Nous avons également simulé  $r = 10\,000$  réalisations de  $\hat{\pi}$  et  $\tilde{\pi}$  pour  $\pi = 0,95$ ,  $L = -0,15$ ,  $U = 0,15$ ,  $\mu = 0$  et  $\mu = 0,10$ ,  $n = 5$ ,  $n = 10$ ,  $n = 15$  et  $n = 20$ . Le Tableau 1 résume les résultats obtenus, avec la convention que, si  $(x_1, \dots, x_r)$  est un  $r$ -échantillon, alors  $\text{Variance} = \frac{1}{r} \cdot \sum_{i=1}^r (x_i - \bar{x})^2$ ,  $\text{MSE} = \text{Variance} + \text{Biais}^2$ ,  $\text{Asymétrie} = \frac{1}{r} \cdot \sum_{i=1}^r (x_i - \bar{x})^3 / \text{Variance}^{3/2}$  et  $\text{Aplatissement} = \frac{1}{r} \cdot \sum_{i=1}^r (x_i - \bar{x})^4 / \text{Variance}^2$ . Les valeurs d'asymétrie et d'aplatissement motivent les développements d'Edgeworth proposés ci-après.

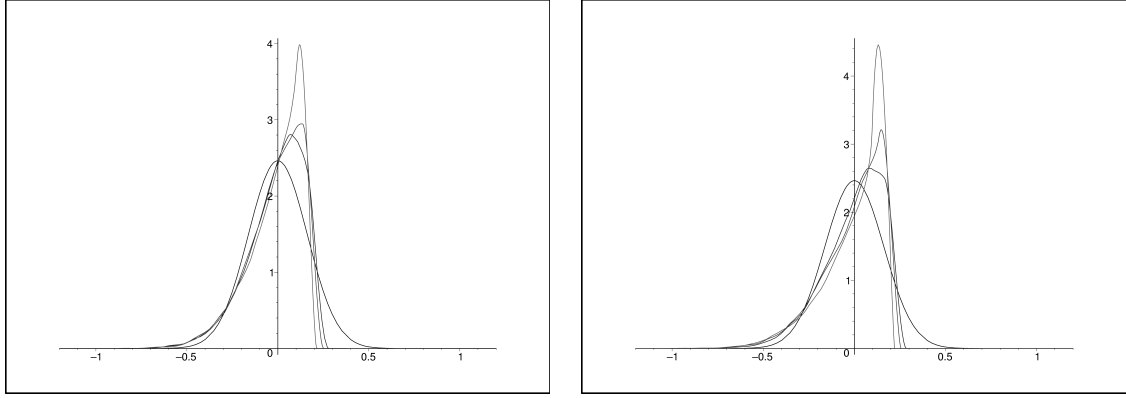


FIG. 1 – Les 2 simulations - voir texte.

$\mu$	0				0, 10			
$n$	5	10	15	20	5	10	15	20
Biais $\hat{\pi}$	-0,0056	-0,0030	-0,0033	-0,0017	0,0016	0,0012	0,0012	0,0001
Variance $\hat{\pi}$	0,0042	0,0022	0,0016	0,0012	0,0038	0,0021	0,0014	0,0011
MSE $\hat{\pi}$	0,0042	0,0022	0,0016	0,0012	0,0038	0,0021	0,0014	0,0011
Asymétrie $\hat{\pi}$	-1,5591	-1,1599	-1,0082	-0,9218	-1,8825	-1,3775	-1,1852	-1,0277
Aplatissement $\hat{\pi}$	5,3367	4,2283	3,9795	3,7984	7,1694	5,0787	4,7367	4,2284
Biais $\tilde{\pi}$	0,0012	0,0001	-0,0011	-0,0006	-0,0011	0,0008	0,0005	0,0001
Variance $\tilde{\pi}$	0,0057	0,0027	0,0018	0,0013	0,0061	0,0026	0,0018	0,0013
MSE $\tilde{\pi}$	0,0057	0,0027	0,0018	0,0013	0,0061	0,0026	0,0018	0,0013
Asymétrie $\tilde{\pi}$	-1,7887	-1,3490	-1,0547	-0,9403	-1,7460	-1,3472	-1,1293	-0,9982
Aplatissement $\tilde{\pi}$	5,8690	4,7935	3,8923	3,8286	5,8653	4,8502	4,2291	3,9678

TAB. 1 – Simulations de  $\hat{\pi}$  et  $\tilde{\pi}$ .

#### 4. Développements d'Edgeworth pour $\hat{\pi}$ et $\tilde{\pi}$

Afin d'obtenir un développement d'Edgeworth pour  $\hat{\pi}$ , nous utilisons les résultats et les notations de [5]. Soient  $\{X, X_1, \dots, X_n\}$  des variables aléatoires indépendantes et identiquement distribuées de loi  $\mathcal{N}(\mu, \sigma^2)$ . Introduisons  $\mathbf{X} = (X, X^2)$  et  $\mathbf{X}_s = (X, X^2, X^3, X^4)$ . Les expressions (2.2) et (2.3) permettent de dire que  $\hat{\pi}$  et  $\tilde{\pi}$  dépendent de  $\bar{X}$  et de  $s_c/k$ . Mais nous pouvons aussi dire que  $\hat{\pi}$  et  $\tilde{\pi}$  sont des fonctions de l'échantillon  $\{(X_i, X_i^2) : 1 \leq i \leq n\}$ .

Maintenant définissons la fonction  $F_{0,MV}$  (resp. la fonction studentisée  $F_{s,MV}$ ) de la façon suivante :

$$F_{0,MV} : (x, y) \in \mathbb{R}^2 \mapsto \widetilde{g}_{MV}(x, y) / \sigma_{MV} \in \mathbb{R} , \quad (4.1)$$

$$F_{s,MV} : (x, y, z, t) \in \mathbb{R}^4 \mapsto \widetilde{g}_{MV}(x, y) / \sqrt{\widehat{\sigma}_{MV}^2(x, y, z, t)} \in \mathbb{R} \quad (4.2)$$

où  $\widehat{\sigma}_{MV}^2(x, y, z, t) = \sum_{i=1}^d \sum_{j=1}^d (\partial \widetilde{g}_{MV} / \partial x_i \cdot \partial \widetilde{g}_{MV} / \partial x_j)(x, y) (\phi(i+j) - \phi(i)\phi(j))$ , avec  $\phi(k)$  la  $k$ -ème composante du vecteur  $(x, y, z, t)$  et  $\widetilde{g}_{MV}(x, y) = g_{MV}(x, y - x^2) - \pi(\mu, \sigma^2)$ .

**Théorème 4.1.** Nous avons le développement d'Edgeworth à l'ordre deux suivant :

$$\mathbb{P}[\sqrt{n}F_{\bullet, MV}(\bar{X}, s_c^2/k^2) \leq x] = \Phi(x) + n^{-1/2}p_{1,\bullet, MV}(x)\phi(x) + n^{-1}p_{2,\bullet, MV}(x)\phi(x) + O(n^{-3/2}), \quad (4.3)$$

uniformément en  $x \in \mathbb{R}$  et où  $p_{1,\bullet,MV}$ ,  $p_{2,\bullet,MV}$  sont définis par les équations (2.24) et (2.25) page 48 dans [5]. Les coefficients  $k_{1,2,MV}$ ,  $k_{2,2,MV}$ ,  $k_{3,1,MV}$  et  $k_{4,1,MV}$  diffèrent d'une statistique à l'autre puisque :

$$\begin{aligned}
k_{1,2,MV} &= \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d a_{ij} \mu_{ij}, & k_{3,1,MV} &= \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d a_i a_j a_k \mu_{ijk} + 3 \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d \sum_{l=1}^d a_i a_j a_k l \mu_{ik} \mu_{jl}, \\
k_{2,2,MV} &= \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d a_i a_j k \mu_{ijk} + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d \sum_{l=1}^d a_{ij} a_{kl} \mu_{ik} \mu_{jl} + \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d \sum_{l=1}^d a_i a_j k l \mu_{ij} \mu_{kl}, \\
k_{4,1,MV} &= \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d \sum_{l=1}^d a_i a_j a_k a_l (\mu_{ijkl} - 3\mu_{ij} \mu_{kl}) + 12 \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d \sum_{l=1}^d \sum_{m=1}^d a_i a_j a_k a_l m \mu_{il} \mu_{jkm} \\
&\quad + \frac{3}{2} \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d \sum_{l=1}^d \sum_{m=1}^d \sum_{o=1}^d a_i a_j a_k l a_{mo} (\mu_{ij} \mu_{kl} \mu_{mo} + 2\mu_{ij} \mu_{km} \mu_{lo} + 4\mu_{ik} \mu_{jl} \mu_{mo} + 16\mu_{ik} \mu_{jm} \mu_{lo}) \\
&\quad + \frac{2}{3} \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d \sum_{l=1}^d \sum_{m=1}^d \sum_{o=1}^d a_i a_j a_k a_l m o (4\mu_{im} \mu_{kl} \mu_{jo} + 2\mu_{im} \mu_{kj} \mu_{lo}),
\end{aligned}$$

où  $a$  est égal à  $F_{0,MV}$  (resp. à  $F_{s,MV}$ ), en désignant par  $^{(i)}$  la  $i$ -ème composante d'un vecteur et avec

$$a_{i_1 \dots i_j} = \left( \partial^j / \partial x^{(i_1)} \dots \partial x^{(i_j)} \right) a(\mathbf{X})|_{\mathbf{X}=\mathbb{E}[\mathbf{X}]} \quad \text{et} \quad \mu_{i_1 \dots i_j} = \mathbb{E} \left[ (\mathbf{X} - \mu)^{(i_1)} \dots (\mathbf{X} - \mu)^{(i_j)} \right].$$

De façon similaire, nous obtenons un résultat pour  $\tilde{\pi}$ . Nous introduisons la fonction  $F_{0,SBVM}$  (ou la version studentisée de cette statistique  $F_{s,SBVM}$ ) définie simplement en remplaçant  $g_{MV}$  par  $g_{SBVM}$  dans les expressions (4.1) et (6). Toutefois un raffinement supplémentaire est nécessaire puisque  $\widehat{g_{SBVM}}(x, y)$  et donc  $F_{0,SBVM}(x, y)$  et  $F_{s,SBVM}(x, y, z, t)$  dépendent de  $n$  (voir [5] page 58 pour plus de détails sur la manière de traiter cette difficulté).

**Théorème 4.2.** Nous avons le développement d'Edgeworth à l'ordre deux suivant :

$$\mathbb{P}[\sqrt{n}F_{\bullet,SBVM}(\bar{X}, s_c^2/k^2) \leq x] = \Phi(x) + n^{-1/2} p_{1,\bullet,SBVM}(x) \phi(x) + n^{-1} p_{2,\bullet,SBVM}(x) \phi(x) + O(n^{-3/2}), \quad (4.4)$$

uniformément en  $x \in \mathbb{R}$  et où  $p_{1,\bullet,SBVM}$  et  $p_{2,\bullet,SBVM}$  ont été définis au Théorème 4.1 et où  $k_{1,2,SBVM}$ ,  $k_{3,1,SBVM}$ ,  $k_{2,2,SBVM}$  et  $k_{4,1,SBVM}$  sont donnés par :

$$\begin{aligned}
k_{1,2,SBVM} &= k_{1,2,MV}, & k_{3,1,SBVM} &= k_{3,1,MV}, & k_{2,2,SBVM} &= k_{2,2,MV} + 2 \sum_{i=1}^d \sum_{j=1}^d b_{0,i} b_{2,j} \mu_{ij}, \\
k_{4,1,SBVM} &= k_{4,1,MV},
\end{aligned}$$

et avec les  $b_{0,i}$ ,  $1 \leq i \leq d$ , et les  $b_{2,j}$ ,  $1 \leq j \leq d$ , définis par l'équation (2.38) page 58 dans [5].

## 5. Développements de Cornish-Fisher

Nous souhaitons obtenir les deux quantiles de niveau  $\alpha$  pour les deux estimateurs dont nous avons obtenu un développement d'Edgeworth qui sont déterminés par :

$$w_{\alpha}^{\bullet,\bullet} = \inf \{ x : \mathbb{P}[\sqrt{n}F_{\bullet,\bullet}(\bar{X}, s_c^2/k^2) \leq x] \geq \alpha \}.$$

Notons  $z_\alpha$  le quantile de niveau  $\alpha$  d'une loi normale centrée réduite. Les équations (4.3) et (4.4) permettent d'obtenir un développement de  $w_\alpha$  dans les quatre cas qui nous intéressent en fonction de  $z_\alpha$  :

$$w_\alpha^{\bullet,\bullet} = z_\alpha - n^{-1/2}p_{1,\bullet,\bullet}(z_\alpha) + n^{-1}p_{1,\bullet,\bullet}(z_\alpha)p'_{1,\bullet,\bullet}(z_\alpha) - \frac{1}{2}z_\alpha p_{1,\bullet,\bullet}^2(z_\alpha) - p_{2,\bullet,\bullet}(z_\alpha) + O\left(n^{-3/2}\right) \quad (5.1)$$

avec  $p_{1,\bullet,\bullet}$  et  $p_{2,\bullet,\bullet}$  comme définis ci-dessus au Théorème 4.1 ou au Théorème 4.2.

## 6. Conclusion et perspectives

Cette Note a permis d'établir les développements d'Edgeworth de deux estimateurs d'une proportion de mesures. Ces développements doivent maintenant être testés sur des données réelles. De tels résultats seront présentés ultérieurement, en particulier le recours à des techniques bootstrap justifié par le fait que les paramètres  $\mu$  et  $\sigma^2$  sont inconnus. Pour plus de détails à ce sujet nous renvoyons à [3] et [4].

## Remerciements

Les auteurs tiennent à remercier les deux rapporteurs anonymes pour leurs suggestions qui ont permis d'améliorer la rédaction de cette Note.

## Références

- [1] F. Bertrand et M. Maumy, Développements d'Edgeworth et intervalles de confiance de deux estimateurs d'une proportion de mesures, *Prépublication de l'IRMA*, (2007).
- [2] T.L. Boullion, G.C. Cascio et J.P. Keating, Comparison of estimators of the fraction defective in the normal distribution, *Communications in Statistics : Theory and Methods*, 14 (1985), 1511-1529.
- [3] A. Guillou, Efficient weighted bootstraps for the mean, *Journal of Statistical Planning and Inference*, 77 (1999), 11-35.
- [4] A. Guillou, Weighted bootstraps for the variance, *Journal of Statistical Planning and Inference*, 81 (1999), 113-120.
- [5] P. Hall, *The Bootstrap and Edgeworth Expansion*, Springer, New York, 1992.
- [6] R. Iasnogorodski et H. Lhéritier, *Théorie de l'estimation ponctuelle paramétrique*, EDP Sciences, Les Ulis, 2003.
- [7] A.N. Kolmogorov, Unbiased estimates, *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya*, 14 (1950), 303-326.
- [8] Maplesoft, Maple 10.0, Waterloo, 2005.
- [9] R.W. Mee, Estimation of the percentage of a normal distribution lying outside a specified interval, *Communications in Statistics : Theory and Methods*, 17 (1988), 1465-1479.