

Une nouvelle approche du sondage aléatoire simple

Martin Körnig
m.kornig@wanadoo.fr

Résumé

Une approche bayésienne du sondage aléatoire simple offre des solutions simples, pratiques et relativement faciles à exploiter numériquement. Il s'agit de solutions analytiques (loi de probabilité, fonction de répartition, espérance, variance) permettant de prédire, pour une classe donnée, le nombre de représentants dans la population.

Mots-clés

sondage aléatoire simple, statistique bayésienne

Abstract

A bayesian approach to statistical surveys using random sampling leads to simple and practical solutions that are relatively easy to numerically implement. The given analytical solutions (probability function, cumulative distribution function, expected value, variance) enable the prediction of the number of individuals in the population that belong to a given class.

Key words

statistical survey, random sampling, bayesian statistics

1. Introduction

Je reprends un problème « classique » : le sondage aléatoire simple (voir par exemple [1]). Sondage où

- les individus sont classés par rapport à un caractère (qualitatif) en plusieurs classes distinctes.
- l'objectif est de prédire, pour chaque classe, le nombre de représentants dans la population en analysant un échantillon aléatoire.

Pour un tel sondage une théorie bayésienne existe [2] (voir également [3] et [4] qui ne traitent que le cas de deux classes). Dans le cas de trois ou davantage de classes, l'application de cette théorie demandait jusqu'ici une intégration numérique importante. Le présent article enlève cette difficulté en donnant des solutions analytiques à la section 5.

Le Tab. 1 rassemble les notations principales adoptées dans cet article.

Pour une classe i quelconque, le nombre n_i (voir Tab. 1) n'est généralement pas connu avec exactitude - ni avant ni après le sondage. Il s'agit d'une variable aléatoire, dont on cherche à spécifier la loi de probabilité (réponse à la section 5.1).

Dans le cas théorique $K = N$, il n'y a pas vraiment de problème de prédiction puisque la population entière a été échantillonnée. C'est le seul cas, d'ailleurs, où les nombres n_1, n_2, \dots, n_I sont connus exactement. L'ensemble de la théorie exposée dans les sections 3 et 5 reste néanmoins valable dans ce cas là.

L'article continue avec une discussion sur l'importance du nombre de classes I (section 6) et une suggestion concernant l'information a priori (section 7).

Tab. 1 : Notations adoptées

Symbole	Domaine	Signification	Contrainte
I	$I \geq 2$	nombre de classes distinctes	
i	$1 \leq i \leq I$	index de classe	
N	$N \geq 1$	taille de la population (nombre d'individus dans la population)	$\sum_{i=1}^I n_i = N$
n_i	$0 \leq n_i \leq N$	nombre présumé de représentants de la classe i dans la population	$\sum_{i=1}^I n_i = N$
n	$0 \leq n \leq N$	nombre présumé de représentants d'une classe (spécifiée ailleurs) dans la population	
K	$0 \leq K \leq N$	taille de l'échantillon (nombre d'individus dans l'échantillon)	$\sum_{i=1}^I k_i = K$
k_i	$0 \leq k_i \leq K$	nombre observé de représentants de la classe i dans l'échantillon	$\sum_{i=1}^I k_i = K$

2. Hypothèse fondamentale

Dès la conception du sondage, on doit connaître le nombre I de classes et la taille N de la population (voir Tab. 1).

Hypothèse 1 : Toutes les répartitions $(n_1, n_2, \dots, n_I) \in \{0, 1, \dots, N\}^I$, satisfaisant $\sum_{i=1}^I n_i = N$, sont initialement équiprobables.

Cette hypothèse reflète la conception de l'expérience aléatoire associée au sondage [2].

3. La loi de probabilité initiale

Avant le sondage, la loi de probabilité associée à une classe i quelconque est nommée p_i^{ini} . p_i^{ini} décrit l'incertitude initiale concernant cette classe. Attention : il ne s'agit pas d'information a priori que l'utilisateur peut choisir ou modifier à volonté (section 7) ; cette loi est une conséquence implicite de l'hypothèse 1. On trouve

$$p_i^{\text{ini}} : \{0, 1, \dots, N\} \rightarrow [0, 1],$$

$$(1) \quad p_i^{\text{ini}}(n) = \frac{C_{N-n+I-2}^{I-2}}{C_{N+I-1}^{I-1}}$$

pour chacune des classes $i = 1, 2, \dots, I$.

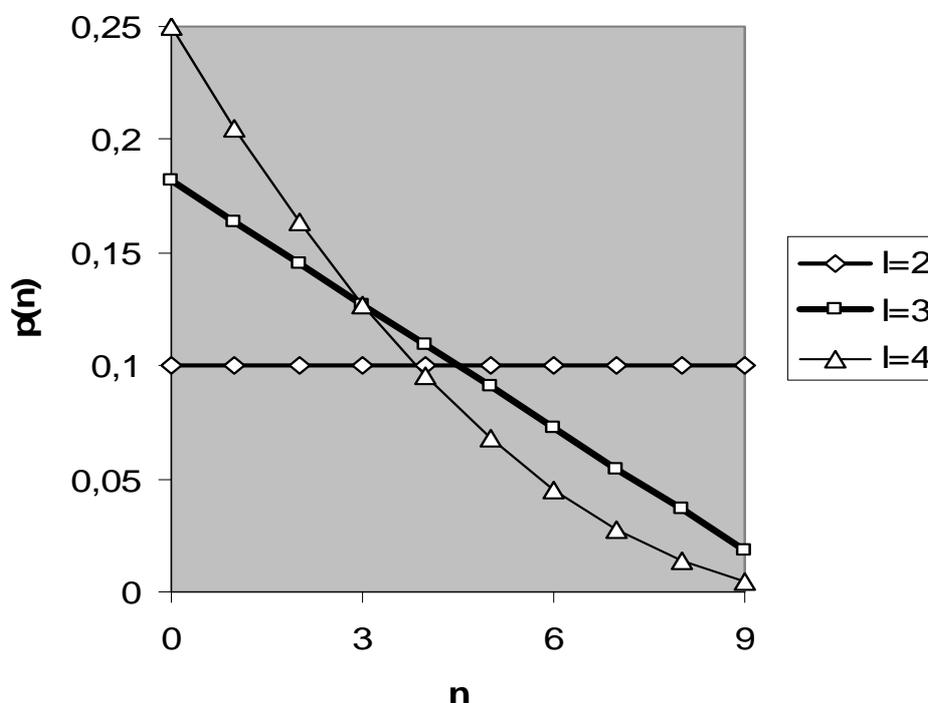
Spécification concernant la notation : Pour deux entiers $a, b \geq 0$,

$$(2) \quad C_a^b = \begin{cases} 0 & \text{si } a < b \\ 1 & \text{si } b = 0 \\ \frac{a(a-1)\cdots(a-b+1)}{1\cdot 2\cdots b} & \text{sinon} \end{cases}$$

dénote le nombre de combinaisons de b éléments choisis parmi a . On notera que la définition (2) admet le cas $a < b$ ce qui sera important pour les équations (4), (5c) et (5d).

Voir Fig. 1 pour quelques exemples de loi (1) pour des sondages de petite taille ($N = 9$). La « forme » principale de p_i^{ini} (uniforme, linéaire, parabolique etc.) dépend de I et ne change pas en variant N .

Fig. 1 : Lois de probabilité initiale p_i^{ini} pour $N = 9$ et trois valeurs différentes de I .



Remarque 1 : L'espérance de n_i est définie par $E_i^{\text{ini}} = \sum_{n=0}^N n p_i^{\text{ini}}(n)$ ce qui donne en fait

$$(3) \quad E_i^{\text{ini}} = \frac{N}{I}.$$

4. Ce que l'on sait à la suite du dépouillement

Après le dépouillement de l'échantillon, on connaît la taille K de l'échantillon et les nombres k_i (voir Tab. 1).

5. Ce que l'on souhaite savoir

Pour chaque classe i , on souhaite connaître :

- la loi de probabilité a posteriori p_i^{post}

- la fonction de répartition a posteriori F_i^{post}
- l'espérance a posteriori E_i^{post}
- la variance a posteriori V_i^{post}

5.1 La loi de probabilité

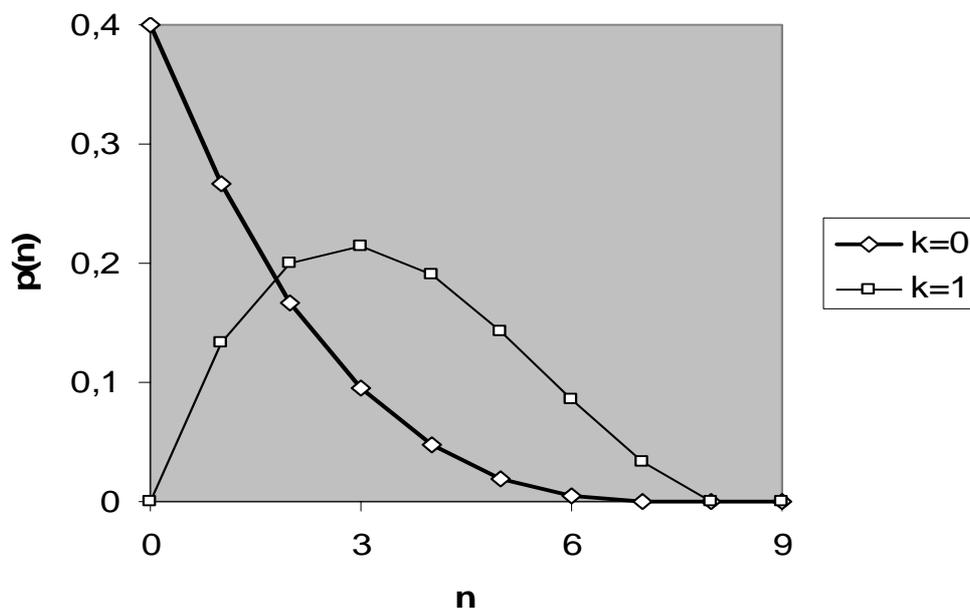
Le résultat principal est la loi de probabilité a posteriori

$$p_i^{\text{post}} : \{0,1,\dots,N\} \rightarrow [0,1],$$

$$(4) \quad p_i^{\text{post}}(n) = \frac{C_n^{k_i} C_{N-n+I-2}^{K-k_i+I-2}}{C_{N+I-1}^{K+I-1}}.$$

Pour $I = 2$, on retrouve cette formule dans certains livres de statistique (par exemple [3], page 147). En général, ces lois sont asymétriques. Elles ressemblent souvent à des gaussiennes (surtout si $N \gg K \gg I$, $K \gg k_i \gg 1$), parfois à des exponentielles ($N \gg 1$, $k_i = 0$), parfois à des paraboles ($N \gg I = K = 2$, $k_i = 1$), parfois à des fonctions de Dirac ($K \approx N \gg 1$). Voir Fig. 2 pour des exemples de visualisations graphiques.

Fig. 2 : Lois de probabilité a posteriori $p_i^{\text{post}}(n)$ pour $N = 9$, $K = 3$, $I = 2$ et deux valeurs différentes de k_i .



Remarque 2 : Grâce à la définition (2), on observe

$$(n < k_i \vee n > N - K + k_i) \Rightarrow p_i^{\text{post}}(n) = 0.$$

5.2 La fonction de répartition

La fonction de répartition a posteriori est définie par

$$F_i^{\text{post}} : \{0,1,\dots,N\} \rightarrow [0,1],$$

$$(5a) \quad F_i^{\text{post}}(n) = \sum_{j=0}^n p_i^{\text{post}}(j).$$

Au lieu d'évaluer $F_i^{\text{post}}(n)$ par (5a), on peut l'évaluer parfois avec beaucoup moins d'opérations arithmétiques à effectuer par l'une des formules suivantes :

$$(5b) \quad F_i^{\text{post}}(n) = \begin{cases} F_i^{\text{post}}(n_0) + \sum_{j=n_0+1}^n p_i^{\text{post}}(j), & \text{si } n_0 < n \\ F_i^{\text{post}}(n_0) - \sum_{j=n+1}^{n_0} p_i^{\text{post}}(j), & \text{sinon} \end{cases}$$

où $n_0 \in \{0, 1, \dots, N\}$ est un point de référence quelconque où la valeur de F_i^{post} est déjà connue,

$$(5c) \quad F_i^{\text{post}}(n) = \frac{\sum_{j=0}^{K-k_i+I-2} C_{N-n+I-2}^j C_{n+1}^{K-j+I-1}}{C_{N+I-1}^{K+I-1}},$$

$$(5d) \quad F_i^{\text{post}}(n) = 1 - \frac{\sum_{j=0}^{k_i} C_{n+1}^j C_{N-n+I-2}^{K-j+I-1}}{C_{N+I-1}^{K+I-1}}.$$

L'efficacité numérique des formules (5) est liée au nombre de combinaisons C_a^b à évaluer. Pour (5a) ce nombre est typiquement de l'ordre de N , bien que pour (5c) et (5d) il est typiquement de l'ordre de K . (5c) et (5d) s'avèrent généralement plus efficaces que (5a) puisque le taux de sondage K/N est souvent faible (typiquement de l'ordre de 10^{-3} ou moindre).

La fonction de répartition F_i^{post} est directement liée à la probabilité associée à un intervalle donné : la probabilité associée à l'intervalle $m_1 < n \leq m_2$, où m_1 et m_2 sont les deux bornes, est $F_i^{\text{post}}(m_2) - F_i^{\text{post}}(m_1)$.

D'autre part, grâce à une stratégie itérative, les formules (5) permettent de calculer un seuil m correspondant à une probabilité P donnée tel que $F_i^{\text{post}}(m) \approx P$. Pour un intervalle de confiance bilatéral à 90%, par exemple, il faut calculer deux bornes m_1 et m_2 tel que $F_i^{\text{post}}(m_1) \approx 0.05$ et $F_i^{\text{post}}(m_2) \approx 0.95$.

5.3 L'espérance

L'espérance a posteriori répond à la définition $E_i^{\text{post}} = \sum_{n=0}^N n p_i^{\text{post}}(n)$ ce qui donne

$$(6) \quad E_i^{\text{post}} = \frac{(1+k_i)(N-K)}{I+K} + k_i.$$

Remarque 3 : $\sum_{i=1}^I E_i^{\text{post}} = N$.

5.4 La variance

La variance a posteriori répond à la définition $V_i^{\text{post}} = \left(\sum_{n=0}^N n^2 p_i^{\text{post}}(n) \right) - (E_i^{\text{post}})^2$. On obtient finalement

$$(7a) \quad V_i^{\text{post}} = \frac{(N-K)(N+I)(1+k_i)(K-k_i+I-1)}{(I+K)^2(I+K+1)}.$$

Remarque 4 : Dans le cas $I = 2$, on notera l'identité $V_2^{\text{post}} = V_1^{\text{post}}$.

En pratique, on remplace souvent la variance par l'écart-type

$$(7b) \quad S_i^{\text{post}} = \sqrt{V_i^{\text{post}}}.$$

et on exprime une prédiction du nombre n_i en employant la notation $E_i^{\text{post}} \pm S_i^{\text{post}}$ (voir exemples au Tab. 2). J'utilise la notation $a \pm b$ pour indiquer l'espérance a et l'écart-type b d'une loi de probabilité quelconque : symétrique ou asymétrique, gaussienne ou non.

Tab. 2 : Prédictions $E_i^{\text{post}} \pm S_i^{\text{post}}$ associées aux lois de probabilité représentées aux Fig. 1 et 2.

Figure	I	N	K	k_i	$E_i^{\text{post}} \pm S_i^{\text{post}}$
Fig. 1	2	9	0	0	4.5±2.9
	3	9	0	0	3.0±2.5
	4	9	0	0	2.1±2.3
Fig. 2	2	9	3	0	1.2±1.3
	2	9	3	1	3.4±1.6

Les formules (4)-(7) restent valables dans le cas $K = 0$ (situation initiale). Dans ce cas, on retrouve alors $p_i^{\text{post}} = p_i^{\text{ini}}$ et $E_i^{\text{post}} = E_i^{\text{ini}}$ (section 3).

6. Le nombre de classes I

6.1 Pourquoi ce nombre est-il essentiel ?

Considérons le cas $K < N$ et une classe i non représentée dans l'échantillon ($k_i = 0$). Puisque l'on ne peut pas exclure la présence d'un (ou plusieurs) représentant(s) de cette classe dans la population non échantillonnée, l'espérance de n_i doit être positive, c'est-à-dire

$$(8) \quad E_i^{\text{post}} > 0.$$

En conservant les valeurs $N, K, k_1, k_2, \dots, k_I$ d'un sondage quelconque, on peut imaginer l'ajout d'une classe $i = I + 1$ non représentée dans l'échantillon ($k_{I+1} = 0$). Afin de satisfaire les relations $\sum_{i=1}^I E_i^{\text{post}} = N$ (avant) et $\sum_{i=1}^{I+1} E_i^{\text{post}} = N$ (après), les espérances $E_1^{\text{post}}, E_2^{\text{post}}, \dots, E_I^{\text{post}}$ ne peuvent pas rester invariables parce qu'il faut « faire de la place » pour la nouvelle classe pour laquelle on accorde $E_{I+1}^{\text{post}} > 0$ selon (8).

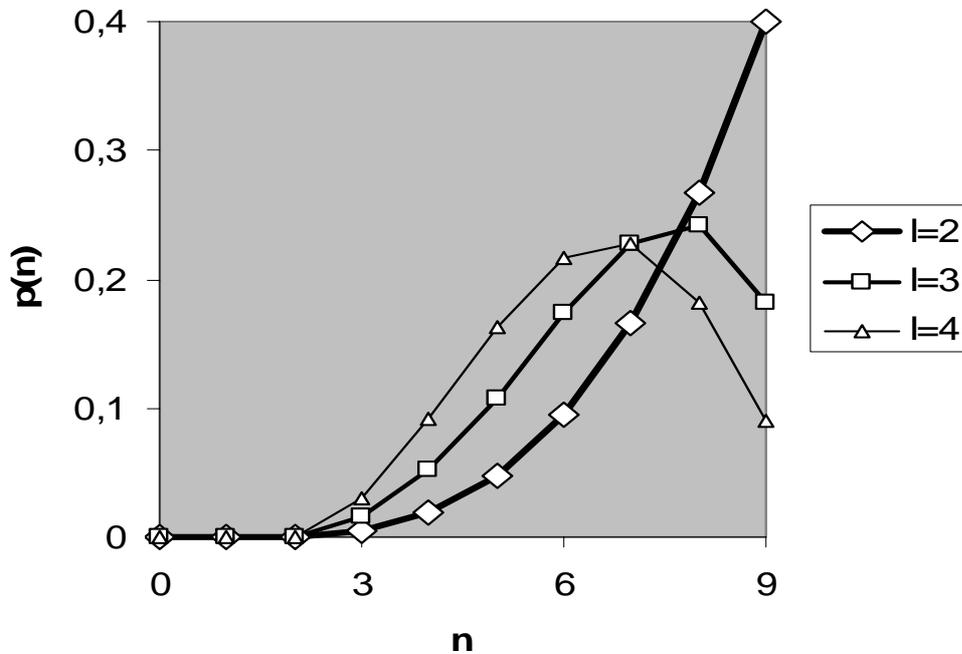
Conclusion : Chaque théorie du sondage aléatoire simple doit prendre en compte le nombre de classes I ; notamment l'espérance E_i^{post} doit dépendre de I .

Bien entendu, la théorie présentée à la section 5 respecte cet impératif. En revanche, une théorie qui extrapole « naïvement » la proportion observée k_i / K à la population entière sans rapport avec I ne peut pas être correcte.

6.2 Comment ce nombre influence-t-il notre prédiction ?

De façon générale : plus I augmente, plus l'espérance E_i^{post} diminue. Ce « comportement » est illustré à la Fig. 3 et au Tab. 3 pour un sondage de petite taille.

Fig. 3 : Lois de probabilité a posteriori $p_i^{\text{post}}(n)$ pour $N = 9$, $k_i = K = 3$ et trois valeurs différentes de I .



Tab. 3 : Prédications $E_i^{\text{post}} \pm S_i^{\text{post}}$ associées aux lois de probabilité représentées à la Fig. 3.

I	N	K	k_i	$E_i^{\text{post}} \pm S_i^{\text{post}}$
2	9	3	3	7.8 ± 1.3
3	9	3	3	7.0 ± 1.5
4	9	3	3	6.4 ± 1.5

Afin d'explorer l'influence de I sur un sondage de taille importante, j'ajoute les Tab. 4 et 5 pour lesquels la taille de la population est $N = 4 \times 10^7$, à peu près le nombre d'adultes en France. Ces deux tableaux donnent

- la proportion estimée $e_i = \frac{E_i^{\text{post}}}{N}$ (en caractère normal),
- l'écart par rapport à la proportion observée $\Delta e_i = \frac{E_i^{\text{post}}}{N} - \frac{k_i}{K}$ (**en gras**) et
- l'écart-type de la proportion estimée $s_i = \frac{S_i^{\text{post}}}{N}$ (*en italique*)

en fonction de I , k_i / K et K , où l'espérance a posteriori E_i^{post} et l'écart-type a posteriori S_i^{post} sont calculées selon les équations (6) et (7).

Tab. 4 : L'influence du nombre de classes, I , sur l'écart Δe_i entre la proportion estimée e_i et la proportion observée k_i / K dans le cas $K = 100$

$\frac{k_i}{K}$	$I = 2$	$I = 5$	$I = 10$	$I = 20$
0	$e_i = 0.0098$ $\Delta e_i = \mathbf{0.0098}$ $s_i = 0.0097$	0.0095 0.0095 <i>0.0094</i>	0.0091 0.0091 <i>0.0090</i>	0.0083 0.0083 <i>0.0083</i>
0.02	0.0294 0.0094 <i>0.0166</i>	0.0286 0.0086 <i>0.0162</i>	0.0273 0.0073 <i>0.0155</i>	0.0250 0.0050 <i>0.0142</i>
0.05	0.0588 0.0088 <i>0.0232</i>	0.0571 0.0071 <i>0.0225</i>	0.0545 0.0045 <i>0.0216</i>	0.0500 0.0000 <i>0.0198</i>
0.1	0.1078 0.0078 <i>0.0306</i>	0.1048 0.0048 <i>0.0297</i>	0.1000 0.0000 <i>0.0285</i>	0.0917 -0.0083 <i>0.0262</i>
0.2	0.2059 0.0059 <i>0.0398</i>	0.2000 0.0000 <i>0.0389</i>	0.1909 -0.0091 <i>0.0373</i>	0.1750 -0.0250 <i>0.0345</i>
0.5	0.5000 0.0000 <i>0.0493</i>	0.4857 -0.0143 <i>0.0485</i>	0.4636 -0.0364 <i>0.0473</i>	0.4250 -0.0750 <i>0.0449</i>
0.8	0.7941 -0.0059 <i>0.0398</i>	0.7714 -0.0286 <i>0.0408</i>	0.7364 -0.0636 <i>0.0418</i>	0.6750 -0.1250 <i>0.0426</i>
0.9	0.8922 -0.0078 <i>0.0306</i>	0.8667 -0.0333 <i>0.0330</i>	0.8273 -0.0727 <i>0.0359</i>	0.7583 -0.1417 <i>0.0389</i>
0.95	0.9412 -0.0088 <i>0.0232</i>	0.9143 -0.0357 <i>0.0272</i>	0.8727 -0.0773 <i>0.0316</i>	0.8000 -0.1500 <i>0.0364</i>
0.98	0.9706 -0.0094 <i>0.0166</i>	0.9429 -0.0371 <i>0.0225</i>	0.9000 -0.0800 <i>0.0285</i>	0.8250 -0.1550 <i>0.0345</i>
1	0.9902 -0.0098 <i>0.0097</i>	0.9619 -0.0381 <i>0.0186</i>	0.9182 -0.0818 <i>0.0260</i>	0.8417 -0.1583 <i>0.0332</i>

Tab. 5 : L'influence du nombre de classes, I , sur l'écart Δe_i entre la proportion estimée e_i et la proportion observée k_i / K dans le cas $K = 1000$

$\frac{k_i}{K}$	$I = 2$	$I = 5$	$I = 10$	$I = 20$
0	$e_i = 0.0010$ $\Delta e_i = \mathbf{0.0010}$ $s_i = 0.0010$	0.0010 0.0010 <i>0.0010</i>	0.0010 0.0010 <i>0.0010</i>	0.0010 0.0010 <i>0.0010</i>
0.02	0.0210 0.0010 <i>0.0045</i>	0.0209 0.0009 <i>0.0045</i>	0.0208 0.0008 <i>0.0045</i>	0.0206 0.0006 <i>0.0044</i>
0.05	0.0509 0.0009 <i>0.0069</i>	0.0507 0.0007 <i>0.0069</i>	0.0505 0.0005 <i>0.0069</i>	0.0500 0.0000 <i>0.0068</i>
0.1	0.1008 0.0008 <i>0.0095</i>	0.1005 0.0005 <i>0.0095</i>	0.1000 0.0000 <i>0.0094</i>	0.0990 -0.0010 <i>0.0093</i>
0.2	0.2006 0.0006 <i>0.0126</i>	0.2000 0.0000 <i>0.0126</i>	0.1990 -0.0010 <i>0.0126</i>	0.1971 -0.0029 <i>0.0124</i>
0.5	0.5000 0.0000 <i>0.0158</i>	0.4985 -0.0015 <i>0.0158</i>	0.4960 -0.0040 <i>0.0157</i>	0.4912 -0.0088 <i>0.0156</i>
0.8	0.7994 -0.0006 <i>0.0126</i>	0.7970 -0.0030 <i>0.0127</i>	0.7931 -0.0069 <i>0.0127</i>	0.7853 -0.0147 <i>0.0129</i>
0.9	0.8992 -0.0008 <i>0.0095</i>	0.8965 -0.0035 <i>0.0096</i>	0.8921 -0.0079 <i>0.0098</i>	0.8833 -0.0167 <i>0.0100</i>
0.95	0.9491 -0.0009 <i>0.0069</i>	0.9463 -0.0037 <i>0.0071</i>	0.9416 -0.0084 <i>0.0074</i>	0.9324 -0.0176 <i>0.0079</i>
0.98	0.9790 -0.0010 <i>0.0045</i>	0.9761 -0.0039 <i>0.0048</i>	0.9713 -0.0087 <i>0.0053</i>	0.9618 -0.0182 <i>0.0060</i>
1	0.9990 -0.0010 <i>0.0010</i>	0.9960 -0.0040 <i>0.0020</i>	0.9911 -0.0089 <i>0.0030</i>	0.9814 -0.0186 <i>0.0042</i>

Dans ces deux tableaux, on notera que la proportion estimée e_i (colonnes 2 à 5, caractère normal) diffère souvent de la proportion observée k_i / K (première colonne). L'écart Δe_i (colonnes 2 à 5, en gras) est souvent négligeable par rapport à l'écart-type de la proportion estimée s_i (colonnes 2 à 5, en italique), c'est-à-dire que l'on a $|\Delta e_i| < s_i$. Cet écart augmente cependant de façon significative, c'est-à-dire $|\Delta e_i| > s_i$, si I et k_i / K sont grands (cases grises). J'en conclus que prendre la proportion observée k_i / K et l'adopter pour la population entière n'est pas toujours une bonne idée. Une telle stratégie peut donc conduire à des erreurs de prédiction significatives si I et k_i / K sont grands que la taille du sondage soit petite ou grande.

7. Comment incorporer une information a priori ?

Considérons d'abord un sondage « en cours » où les données arrivent l'une après l'autre. On peut s'arrêter à un moment quelconque et analyser/visualiser ce que l'on a déjà. A ce moment, on peut considérer que les données déjà prises en compte fournissent l'information a priori pour celles qui vont arriver.

Dans la même optique, afin de créer une information a priori, je propose de récupérer les données d'un sondage antérieur du même type k_i^{prior} et $K^{\text{prior}} = \sum_{i=1}^I k_i^{\text{prior}}$ (ou d'« inventer » les données d'un pré-sondage fictif) et de les combiner avec celles du sondage actuel k_i^{act} et $K^{\text{act}} = \sum_{i=1}^I k_i^{\text{act}}$ (ou réel) en additionnant

$$(9a) \quad k_i = k_i^{\text{prior}} + k_i^{\text{act}},$$

$$(9b) \quad K = K^{\text{prior}} + K^{\text{act}}.$$

Cette méthode est souple et facile à mettre en œuvre. De plus, on est sûr de commencer avec une information a priori cohérente.

8. Conclusion

Pour exprimer une prédiction simple du nombre de représentants n_i d'une classe i après un sondage aléatoire simple, je préconise la notation $E_i^{\text{post}} \pm S_i^{\text{post}}$, où E_i^{post} signifie l'espérance et S_i^{post} l'écart-type. C'est facile à calculer (formules (6) et (7)). C'est facile à communiquer. Et c'est très souvent suffisant. Les paramètres E_i^{post} et S_i^{post} sont définis pour une loi de probabilité quelconque, qu'il s'agisse d'une gaussienne ou non.

Pour ceux qui ont besoin d'une précision supérieure, je donne également les formules pour évaluer la loi de probabilité (section 5.1) et la fonction de répartition (section 5.2).

La théorie est valable sans rapport avec les tailles N et K et met en évidence l'importance du nombre de classes I . Ces résultats ont un champ d'application intéressant car ils offrent les possibilités suivantes :

- Tracer les lois de probabilité a posteriori (voir Fig. 2 et 3). Grâce à un logiciel spécifique évitant les *overflows* et à la puissance des ordinateurs personnels d'aujourd'hui, ceci est faisable même pour un sondage de taille importante. Ainsi, la communication des résultats d'un sondage pourrait être accompagnée d'un (ou plusieurs) graphique(s) intuitif(s).
- Traiter un sondage de « petite taille ».
- Traiter un sondage avec un nombre élevé de classes.

Une théorie « simple » qui prend la proportion observée k_i / K et l'adopte directement pour la population entière (sans prendre en compte le nombre de classes I) encourt des erreurs de prédiction significatives si I et k_i / K sont grands (section 6).

Dans le cas $I = 2$, les prédictions de la nouvelle approche et celles de l'approche habituelle (voir par exemple [1]) semblent converger de façon asymptotique d'après mes premières expériences numériques (Tab. 4 et 5).

Remerciements

Je remercie Gilles Sado pour des discussions et pour ses conseils concernant les notations et certaines formulations. Qu'Harry-Pascal Bannais soit remercié pour son soutien financier.

Références

- [1] Dussaix, A.-M. & Grosbras, J.-M., 1993. *Les sondages : principes et méthodes*. Presses universitaires de France, série « Que sais-je ? ».
- [2] Körnig, M., 2003. Théorie du sondage aléatoire et étude d'un sondage d'opinion avant le 1er tour d'une élection présidentielle. *La Revue de MODULAD*, 30. 12-18.
- [3] Robert, C., 1992. *L'analyse statistique bayésienne*. Economica, Paris.
- [4] Simar, L., 2002. Le paradigme bayésien. Dans *Méthodes bayésiennes en statistique*, éditeurs : Dreesbeke, J.-J., Fine, J. et Saporta, G., Editions Technip (page 19-60).