Mathématiques et sciences humaines
Mathematics and social sciences

183 | Automne 2008
Hommage à Georges-Th. Guilbaud

A propos de l'estimation de nombre des coins

About the estimated number of coining

Georges-Théodule Guilbaud

Édition électronique
URL : http://journals.openedition.org/msh/10823
DOI : 10.4000/msh.10823
ISSN : 1950-6821

Éditeur
Centre d'analyse et de mathématique sociales de l'EHESS

Édition imprimée
Date de publication : 14 décembre 2008
ISSN : 0987-6936

Référence électronique
Georges-Théodule Guilbaud, « A propos de l'estimation de nombre des coins », *Mathématiques et sciences humaines* [En ligne], 183 | Automne 2008, mis en ligne le 15 décembre 2008, consulté le 20 avril 2019. URL : http://journals.openedition.org/msh/10823 ; DOI : 10.4000/msh.10823

Math. Sci. hum. / Mathematics and Social Sciences (46e année, n° 183, *2008*(3), p. 71-106)

DEUX ARTICLES SCIENTIFIQUES DE G.-Th. GUILBAUD

Il semble naturel de compléter ce numéro de *Mathématiques et Sciences humaines* par des écrits de Guilbaud lui-même.

Nous en avons choisi deux, parus l'un en 1970, l'autre en 1974.

Il y a trois raisons à ces choix :

- Ces deux textes sont aujourd'hui quasiment introuvables.
- Ensuite, ils témoignent de la diversité des intérêts de Guilbaud dans les sciences sociales. Les mathématiques utilisées sont elles aussi très éloignées l'une et l'autre, algèbre dans le premier cas, statistique inférentielle dans le second.
- Enfin, et surtout, ils sont très caractéristiques de la manière dont il présentait idées et techniques mathématiques pour des lecteurs étrangers à cette discipline.

Le premier de ces articles[1] (cf. p. 73-96) fut publié à la suite d'un exposé oral sur des travaux de l'ethnologue océaniste Jean Guiart, dans un domaine qui a toujours intéressé Claude Levi-Strauss, celui des systèmes de parenté.

Le second[2] (cf. p. 97-106) résulte de la très féconde collaboration qui s'était instaurée, dans les années 1970, entre des historiens de la numismatique (et au premier chef, Julien Guey) et le Centre de mathématique sociale, dont la statisticienne Charlotte Carcassonne, et Guilbaud lui-même.

Nous remercions les directions des revues concernées d'avoir autorisé la reproduction de ces textes.

---

[1] « Système parental et matrimonial au Nord Ambrym », *Journal de la Société des Océanistes* n° 26, tome XXVI, mars 1970, p. 9-32.

[2] « À propos de l'estimation du nombre de coins », *Bulletin de la Société française de Numismatique*, Cabinet des médailles de la bibliothèque nationale, 29e année, n° 7 mensuel, juillet 1974, p. 625-634.

GUILBAUD (G.Th.). — **A propos de l'estimation du nombre des coins.**

AVERTISSEMENT

Quelques lectures (*) m'ont induit à plaider pour une réflexion statistique.

...« sous réserve d'une justification théorique dont on ne saurait ici masquer l'absence... »

(GIARD, p. 97.)

...« la formule avancée sans autre forme de procès, par I.D. BROWN... »

(id., p. 96.)

« ...la méthode n'est pas clairement justifiée... »

(id., p. 167.)

...« the validity of the formula itself I am not competent to discuss since I share with Dr. Metcalf the disadvantage of not being a statistician... »

(GRIERSON, p. 153.)

Voilà un échantillon de phrases incidentes qui me paraissent révéler un fâcheux état d'esprit, une renonciation dangereuse. Pour parler du premier consulat d'Octave, M. Giard n'a pas besoin d'être consul, ni d'être Octave. Or la statistique aussi est un phénomène social et historique. Mais la numismatique n'est pas privilégiée : les diverses études scientifiques concernant l'homme, ses sociétés et leur histoire, utilisent toutes des méthodes statistiques. Il n'y a pas de statistique sans mathématique ; mais ce n'est pas une raison pour transformer les procédures et formules en une sorte de sorcellerie qu'on craint, qu'on révère ou qu'on méprise.

Il faut chercher à comprendre.

LE PROBLEME

On a un lot de pièces provenant toutes d'un même monnayage ; on compare ces pièces deux à deux, et l'on peut affirmer que certaines paires (mais non pas toutes) ont été sûrement frappées avec le même coin (1).

Selon que ces « liaisons par le coin » sont plus ou moins nombreuses, on cherche à tirer d'un tel examen quelque information concernant le nombre total des coins utilisés pour frapper cette émission (2).

STATISTIQUE DESCRIPTIVE

La « Com-Paraison », comme son nom l'indique (3), portera sur les *paires :* dans un examen minutieux (ou : prudent) elle devra porter sur *toutes* les paires.

(*) Les renvois sont faits dans le texte par le seul nom de l'Auteur : on trouvera les références bibliographiques à la fin du texte.

Par exemple, si j'ai cinq pièces : A, B, C, D et E, je devrai effectuer les comparaisons suivantes :

A et B			
A et C	B et C		
A et D	B et D	C et D	
A et E	B et E	C et E	D et E

ce qui fait, au total :

$$4 + 3 + 2 + 1 = 10 \text{ comparaisons.}$$

S'il y avait cent pièces, le nombre des comparaisons serait :

$$99 + 98 + 97 + 96 + \dots + 3 + 2 + 1 = 4.950.$$

Au lieu d'additions fastidieuses, on sait, depuis la plus haute antiquité, remplacer le calcul précédent par un plus simple :

$$1 + 2 + 3 + 4 + \dots + (y\text{-}3) + (y\text{-}2) + (y\text{-}1) = y(y\text{-}1)/2$$

Le résultat (4) est désigné par la notation classique $\binom{y}{2}$:

$$\binom{5}{2} = 10,\ \binom{6}{2} = 15,\ \binom{10}{2} = 45,\ \binom{100}{2} = 4.950, \text{ etc.}$$

Ayant effectué $\binom{y}{2}$ examens, c'est-à-dire ayant posé autant de questions du genre :

« Est-ce le même coin qui a frappé ces deux pièces ? »
on aura noté le nombre de réponses positives :

« Oui, c'est le même coin. »

De sorte qu'un premier *résumé statistique* de l'examen pourra se présenter sous la forme :

Ayant examiné un lot de 47 pièces (5), on a dû faire :

$$47 \times 46 /2 = 1.081 \text{ comparaisons}$$

dont 179 ont conclu : même coin de revers.

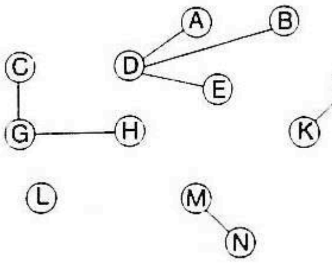
(Notons, pour préparer ce qui va suivre : à peu près un examen sur six a donné une réponse positive.)

La procédure qui vient d'être décrite a été qualifiée de minutieuse (et de : prudente). On peut, en effet, imaginer une autre procédure qu'on pourra dire « rapide » (ou : « économique »).

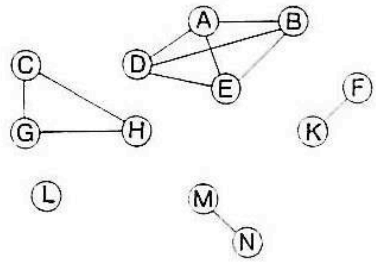
On compare la pièce A et la pièce B : elles ont même coin. On compare ensuite B et C : même coin. Alors, dira-t-on, inutile de confronter A et C : la relation « même coin » est transitive.

On cherche donc à effectuer le nombre minimum de comparaisons.

Dans les dessins ci-dessous, les pièces sont censées posées sur la table, et liées par un trait si le coin est le même. Dans les deux schémas, il s'agit du même lot de pièces.



(rapide)



(minutieuse)

On voit que les expressions, apparemment claires, qu'on peut lire dans les comptes rendus : « le nombre de paires de pièces ayant même coin », ou : « duplications of dies », ou : « linkage », sont ambiguës. Pour le même lot de douze pièces figuré ci-dessus, ce peut être 7 ou 11. Et même, si l'on voulait comprendre littéralement : 2 vraies paires seulement !

Pour y voir clair, instituons une troisième manière de faire, plus méthodique : imaginons qu'on range les pièces dans des boîtes, en plaçant dans une même boîte toutes les pièces de même coin (6).

Puis on range les boîtes selon le nombre des pièces qu'elles contiennent :

z_1 = nombre des pièces solitaires (ou Hapax, ou Singletons)

z_2 = nombre des paires (boîtes à deux pièces seulement)

z_3 = nombre des boîtes qui contiennent trois pièces (on peut dire : triplets, ou : Brelans - (7)

et d'une façon générale pour tout nombre naturel h,

z_h = nombre des coins représentés par h pièces dans le lot.

La statistique complète sera donc représentée par la liste :

$$z_1,\ z_2,\ z_3,\ z_4,\ z_5,\ \ldots$$

Pour le cas fictif dessiné plus haut :

$$z_1 = 1,\ z_2 = 2,\ z_3 = 1,\ z_4 = 1,\ z_5 = 0.$$

Pour les Aurei d'Octave on a, selon la source citée :

$h =$	1	2	3	4	5	6	7	8	9	10	11	12	13
$z_h =$	3	1	0	0	1	0	1	0	0	0	0	2	0

On pourra souhaiter, selon les besoins, *résumer* une telle statistique.

On peut le faire de diverses manières, en combinant les z_h.

Le *nombre des pièces* du lot est :

$$y = z_1 + 2\,z_2 + 3\,z_3 + 4\,z_4 + \ldots + h\,z_h + \ldots$$

L'expression : « *le nombre des paires* » pourra désigner, selon le contexte :

(stricto sensu) : z_2

(procédure minutieuse) : $z^* = z_2 + 3\,z_3 + 6\,z_4 + 10\,z_5 +$ etc.

(procédure rapide) : $z^{**} = z_2 + 2\,z_3 + 3\,z_4 + 4\,z_5 +$ etc.

On peut encore considérer comme digne d'intérêt le résumé suivant :

$$z_1 + z_2 + z_3 + z_4 + \text{etc.}$$

qui n'est autre que :

le *nombre des coins* représentés dans le lot

et qui se calcule comme : y — z**.

Toutes les formules sont du même type : somme de tous les z_h préalablement multipliés par des coefficients convenables (8).

On dit aussi : somme pondérée, ou : combinaison linéaire.

LA REGLE DE BROWN

« (loc. cit. p. 580) ... The number of occasions on which two coins « struck from the same die appear in a find provides a means of « estimating the number of dies used...

« (footnote) : It can be shown that if *x* dies were used in pre- « paring a coinage, in a sample of *y* coins, *z* pairs of coins would be « expected to be struck from the same die, where :

$$x = y\,(y\text{-}1)\,/\,2\,z$$

« provided that *z* is small compared with *y*.

« For example, in a sample of 100 coins, three pairs of coins « were found from the same obverse die, therefore

« no. of dies = 99 × 100 / 2 × 3 = 1.650.

Par sa concision même, ce texte risque fort (9) d'induire en erreurs, bien qu'à strictement parler il n'en contienne pas.

Le lecteur pressé ne retiendra qu'une formule en *x*, y et z.

Mais que signifient ces trois lettres ?

1° *x* est présenté comme le nombre des coins employés dans l'émission. Or, ce nombre est inconnu : ce serait instituer quelque arithmomancie que de prétendre qu'une formule, quelle qu'elle soit, permet de deviner le nombre des coins.

Il vaudrait mieux dire : ce que donne la formule sous le nom de *x* c'est une

ESTIMATION

Pour ma part, j'aurais préféré écrire quelque chose (10) comme ceci :

$$\text{Estim}(x) = y\,(y\text{-}1)\,/\,2\,z.$$

2° y désigne le nombre des pièces dans le lot étudié : pas de difficulté, il suffit de compter (et recompter, pour être sûr). On notera que y n'intervient dans le calcul que par :

$$y\,(y\text{-}1)\,/\,2$$

c'est-à-dire le nombre triangulaire dont il a déjà été question :

$$\binom{y}{2} = \text{nombre de façons de prendre les pièces deux à deux.}$$

3° z présente quelque ambiguïté, comme on l'a dit. Selon le texte : « the number of occasions »... », semble vouloir dire : « le nombre des cas où apparaissent dans le lot deux pièces frappées au même coin », ce qui conduirait à prendre l'interprétation ci-dessus notée par z^*.

En admettant provisoirement cette façon de voir, la Règle préconisée par BROWN est fort simple :

pour examiner un lot de 100 pièces, il faut effectuer :

$$100 \times 99 / 2 = 4.950 \text{ comparaisons ;}$$

on compte le nombre z^* des réponses positives (même coin)

et l'on calcule le rapport : $4.950/z^*$

qui est une estimation de x.

Ce nombre z^* peut, à priori, varier entre deux extrêmes :

il pourrait se faire que $z^* = 0$, si deux pièces du lot ne montraient jamais le même coin. Dans ce cas, la règle est inapplicable.

A l'autre extrême : toutes les pièces du lot sont frappées du même coin ; dans ce cas $z^* = \binom{y}{2}$, nombre maximum et la règle donne : $\text{Estim}(x) = 1$.

Ce qui n'est pas absurde, comme le reconnaît GRIERSON :

« if a random sample of 100 coins, say, is examined ... if all the coins are struck by the same pair of dies, it is overwhelmingly probable that these were the only ones used... »

Nous proposons donc l'interprétation :

$$\text{Estim}(x) = \frac{\binom{y}{2}}{z^*}$$

ce qui signifie :

$$\text{Estim}(x) = \frac{\text{nombre maximum de comparaisons}}{\text{nombre de réponses positives}}$$

Mais : est-ce bien l'intention de l'Auteur ? et surtout : est-ce une bonne règle ? et enfin : par quels raisonnements est-on arrivé à prescrire cette règle ?

LE MODELE DE BROWN

L'intention de l'Auteur, et sa méthode, sont nettement indiquées dans son texte, mais ... par un seul mot !

Et, malheureusement, pour qui n'est pas habitué au dialecte des statisticiens anglo-saxons, ce petit mot risque de passer inaperçu :

« z pairs of coins would be *expected*... »

Bien sûr, « to expect » est un verbe de la langue courante (11) mais dans un contexte statisticien il renvoie au sens technique du substantif « mathematical expectation », lequel correspond sans équivoque au français « espérance mathématique » (12).

Il ne faut donc pas traduire par : « on peut s'attendre à voir z paires », trop anodin en français (13), mais bien par « z étant l'espérance mathématique du nombre de paires ».

La notion d'Espérance Mathématique est la clef de voûte de tout le calcul des probabilités (14). Il ne faut pas dissimuler : tout repose sur les probabilités. Tâchons d'être bref.

Le *modèle probabiliste*, comme on dit, — seul fondement rationnel possible pour une règle d'estimation telle que celle de BROWN — est, dans ses grandes lignes, assez simple, mais lourd de conséquences.

On considère d'abord, par la pensée, la totalité de l'émission : les pièces qui la constituent sont rangées en classes, selon le coin (15) qui les a frappées. On désignera par x le nombre des classes, c'est-à-dire le nombre des coins. On suppose que les diverses classes contiennent le même nombre de pièces ou, au moins, à peu près le même nombre de pièces (16).

On choisit ensuite un nombre entier y qui sera le nombre de pièces du *lot* examiné : on dira d'ailleurs plutôt que y est l'effectif de l'échantillon.

Puis on imagine l'ensemble de *tous* les échantillons *possibles*, prélevés dans l'émission et comportant exactement y pièces. Evidemment, cet ensemble d'échantillons possibles est immense, mais il ne s'agit pas d'en faire le catalogue détaillé !

A chaque échantillon possible correspond une valeur bien déterminée de la caractéristique z (nombre de liens par le coin) : ici, nous prendrons la définition z^*.

On suppose enfin que tous les échantillons possibles sont également probables (17).

Alors le nombre z^*, qui peut varier d'un échantillon à l'autre, devient une « variable aléatoire » ou mieux : un « aléa » numérique : chacune de ses valeurs possibles est dotée d'une probabilité.

Dans le cas qui nous occupe l'*Espérance de l'Aléa* z^* est tout simplement la moyenne (équipondérée) de tous les z^* (une valeur par échantillon). Ou bien, si l'on préfère : la moyenne des valeurs *possibles* pour z^*, pondérées par le nombre d'échantillons fournissant chaque valeur de z^*.

Il ne reste plus qu'à calculer.

UN CALCUL D'ESPERANCE

Il ne peut être question de développer ici le détail des calculs : on se contentera d'une illustration destinée à dissiper toute apparence de mystère et à bien situer le rôle de la technique. Pour pouvoir expliciter les détails sans trop d'algèbre, nous prendrons d'abord des nombres si petits que l'exemple en deviendrait presque ridicule pour qui prendrait les choses au pied de la lettre : mais il ne s'agit que de faire comprendre la méthode.

Voici une émission de trente pièces : trois coins ($x = 3$), chacun frappant dix pièces :

coin A : A0, A1, A2, A3, A4, A5, A6, A7, A8, A9 ;
coin B : B0, B1, B2, B3, B4, B5, B6, B7, B8, B9 ;
coin C : C0, C1, C2, C3, C4, C5, C6, C7, C8, C9.

Catalogue résumé des échantillons de quatre pièces ($y = 4$) :

Premier type : toutes du même coin

Trois variétés : AAAA, BBBB et CCCC.
Chaque variété : 210 échantillons possibles (18).

Deuxième type : un triplet et un hapax

Six variétés : AAAB, AAAC, BBBA, BBBC, CCCA et CCCB.
Chaque variété : 1.200 échantillons.

Troisième type : deux paires

Trois variétés : AABB, AACC et BBCC.
Chaque variété : 2.025 échantillons.

Quatrième type : une seule paire

Trois variétés : AABC, ABBC et ABCC.
Chaque variété : 4.500 échantillons.

Récapitulation (statistique du possible) :

Types	z_1	z_2	z_3	z_4	z^*	Nombre d'échantillons
I	0	0	0	1	6	2 × 210 = 630
					3	6 × 1.200 = 7.200
II	1	0	1	0	2	3 × 2.025 = 6.075
III	0	2	0	0	1	3 × 4.500 = 13.500
IV	2	1	0	0		
						Total = 27.405

Il reste à calculer la moyenne des valeurs de z :

Vaeur de z^* :	Nombre de fois :	
6	630	3.780
3	7.200	21.600
2	6.075	12.150
1	13.500	13.500
	27.405	51.030

Moyenne (ou Espérance) $\dfrac{51.030}{27.405} = \dfrac{54}{29} = 1,862$

On voit le genre de calcul : il suffira de se donner x et y, ainsi que le nombre, que nous désignerons par n, des pièces frappées par chaque coin (ci-dessus on avait pris n = 10).

Il se trouve, à cause de la simplicité du modèle, qu'on peut établir sans peine une formule valable dans tous les cas (19) :

$$\text{Esp.}\ (z^*) = \frac{y(y-1)\,/\,2}{x + \dfrac{x-1}{n-1}}$$

UNE APPROXIMATION

La formule est valable pour toutes valeurs de x, y et n. Mais on n'aura probablement pas l'occasion de s'en servir pour des valeurs invraisemblables. On va voir qu'on peut alors la simplifier, au prix d'une très légère erreur (20).

La formule donne l'Espérance de z comme un quotient :

1° Au numérateur, on reconnaît le nombre triangulaire :

$$\binom{y}{2} = y(y-1)\,/\,2$$

(nombre maximum de paires à examiner dans l'échantillon).

2° Au dénominateur, somme de x (nombre de coins) et d'une fraction. Je dis que, la plupart du temps, cette fraction est assez petite pour qu'on puisse la négliger sans dommage : il sera assez rare, en effet, que x ne soit pas plus petit que n, donc, le plus souvent, ladite fraction est inférieure à l'unité :

x	=	100	100	50	50	50	50
n	=	1.000	10.000	5.000	500	100	50
x-1 / n-1	=	0,0991	0,0099	0,0098	0,098	0,495	1,000

On arrondira donc, en prenant tout simplement x au dénominateur :

$$\text{Approxim. Esp. } (z^*) = \frac{\binom{y}{2}}{x}$$

(conditions de validité : le rapport x/n est négligeable devant x).

C'est finalement cette formule approchée qui a servi de point de départ à BROWN.

UNE METHODE D'ESTIMATION

Dans la description du modèle précédent, les nombres x et y sont des données (on a évacué n qui ne jouait guère). Le nombre z peut varier, même lorsque x et y sont fixés. Grâce au schéma probabiliste (tous les échantillons sont également probables), ce nombre variable devient un *Aléa* (ou : variable aléatoire) dont on peut calculer l'*Espérance*, qui est cette fois fonction de x et de y :

$$\text{Esp. } (z^*) = \frac{\binom{y}{2}}{x}$$

Mais dans la réalité, si y (nombre de pièces dans l'échantillon) est connu, x (le nombre des coins) est inconnu. On va alors essayer diverses hypothèses concernant cette inconnue.

Prenons, par exemple : y = 100 ; d'où : $\binom{y}{2}$ = 4.950.

si : x	=	1.000	1.001	1.002	...	1.010	1.020	...	1.050	..	1.100 ..
alors : Esp. (z^*)	=	4,950	4,945	4,940	...	4,901	4,853	...	4,714	..	4,500 ..

1.500	1.550	1.600	1.650	1.700	1.750	1.800	...	2.000	...
3,300	3,194	3,094	3,000	2,912	2,828	2,750	...	2,475	...

On remarquera ceci : d'assez grandes variations de l'hypothèse x ne changent pas beaucoup le résultat.

La règle de BROWN peut se traduire ainsi : choisissez un x *(hypothétique)* tel que *l'Espérance* de l'Aléa z^* soit aussi voisine que possible de la valeur particulière de z^* que vous aurez effectivement *observée* dans votre enquête.

Ce qu'on peut écrire :

$$\text{Est}\,(x) = \frac{\binom{y}{2}}{\text{Obs}\,(z^*)}$$

et qui consiste, en somme, à lire la table numérique en sens inverse :

$$3{,}000 \longrightarrow 1.650.$$

On admettra aisément qu'il conviendrait d'accompagner cette règle d'un avertissement tel que : « ne soyez pas trop précis ».

Car il faut remarquer encore ceci (nous conservons toujours l'exemple $y = 100$) :

si : obs (z) =	1	2	3	4	5	6	7	8	...
alors : Est (x) =	4.950	2.475	1.650	1.237	990	825	707	619	...

De sorte qu'en exécutant à la lettre la règle d'estimation, on serait conduit à penser que le nombre de coins peut être de 1.650 ou de 1.237 (selon que l'observation aura donné 3 ou 4) mais jamais de valeurs intermédiaires. Ce qui serait inquiétant.

On voit ce qui manque à la règle : une indication formelle de l'incertitude. On dira évidemment : « environ », « à peu près », « grosso modo », « c'est un ordre de grandeur », et ainsi de suite. Mais cela peut ne pas suffire.

C'est cette lacune que LYON a essayé de combler (21).

Pour le moment (22), concluons :

1° La règle de BROWN est commode : elle consiste à calculer le rapport entre deux nombres faciles à observer : ceux que nous avons désignés par :

$$\binom{y}{2} \text{ et } z$$

2° Ne pas se tromper de z (il y en a plusieurs : ici c'est z^*).

3° Ne pas chercher autre chose qu'un « ordre de grandeur raisonnable ».

4° Ne jamais perdre de vue les hypothèses du modèle :

a. Chaque coin a frappé le même nombre n de pièces.

b. Ce nombre est grand (devant x).

c. Tous les échantillons possibles sont également probables.

---

(1) Soit le droit, soit le revers soit les deux (remarque valable dans tout ce qui suit, et qui ne sera pas répétée).

(2) L'intérêt de cette recherche étant d'en pouvoir tirer quelque indication sur le volume de l'émission, en supposant, par exemple, que chaque coin a frappé le même nombre de milliers de pièces.

(3) Certains statisticiens parlent de « comparaisons par paires », n'est-ce pas un pléonasme ?

(4) Qu'on appelle « nombre triangulaire », ou bien « nombre des combinaisons deux à deux de y objets », ou encore « nombre de façons de choisir deux parmi y ».

Certains auteurs français ont préféré naguère la notation C_y^2.

(5) Aurei d'Octave, frappés vers l'an 43 av. J.C. (selon Giard).

(6) Précisons : le rangement achevé, si deux pièces ne sont pas dans la même boîte, c'est qu'elles n'appartiennent pas au même coin.

(7) Giard propose de dire : Trois-Couples, etc. h-Couples. Un tel néologisme serait mal reçu par le vocabulaire mathématique français actuel : « couple » s'y oppose à « paire » d'une part, et, d'autre part, les deux impliquent le binaire.

(8) Pour la combinaison y, les coefficients sont les naturels :

$1, 2, 3, \ldots, h, \ldots$

Pour z*, ce sont les triangulaires :

$0, 1, 3, 6, 10, 15, \ldots, \binom{h}{2}, \ldots$

Pour z**, les naturels encore, mais commençant au zéro :

$0, 1, 2, 3, 4, \ldots, (h-1), \ldots$

(9) Une enquête mais fort sommaire, me permet maintenant de dire : la concision de ce texte absout les erreurs de plusieurs numismates.

(10) Ou bien, comme dans les manuels de Statistique : $\bar{x} = \ldots$

(11) Tous les écoliers savent réciter le fameux Signal de Nelson à Trafalgar.

(12) Et cette « Espérance » a bien peu de rapports avec la Deuxième Vertu.

(13) « On peut espérer voir z paires » est encore plus fallacieux.

(14) Son origine remonte à Pascal lui-même (1654) qui disait : « droit d'attente ». Huygens a introduit le latin « expectatio », immédiatement retraduit (est-ce Leibniz ?) par le français « espérance ».

(15) Le coin droit ou bien le revers : ce sera le même raisonnement pour l'un ou l'autre. Il faudrait quelques modifications, mais peu graves, si l'on tenait compte des deux coins à la fois.

(16) Disons, pour fixer les idées, quelques milliers de pièces par coin. Mais, ici, encore, l'important est de faire des hypothèses assez précises.

(17) C'est le sens le plus habituel de 'random sample'. Ne pas traduire : « pris au hasard », beaucoup trop vague, ou trop précis !

(18) Je ne tiens pas à ennuyer le lecteur avec les techniques classiques de l'énumération combinatoire. Voici cependant comment on a calculé :

210 = 10 × 9 × 8 × 7 / 1 × 2 × 3 × 4
1.200 = 10 × 9 × 8 × 10 / 1 × 2 × 3 × 1
2.025 = 10 × 9 × 10 × 9 / 1 × 2 × 1 × 2
4.500 = 10 × 9 × 10 × 10 / 1 × 2 × 1 × 1
27.405 = 30 × 29 × 28 × 27 / 1 × 2 × 3 × 4

(19) Vérifiez pour notre exemple numérique : x = 3, y = 4, n = 10 ; on trouve Esp (z) = 6 / (3 + 2/9) = 54 / 29.

(20) Dans le jargon usel, on dira qu'on suppose n « infiniment grand » (ce qui est évidemment exagéré !) ou bien qu'on remplace la distribution hypergéométrique par une approximation multinomiale.

(21) En même temps qu'il modifiait la base de l'estimation : au lieu de l'aléa z* il a choisi celui que nous avons appelé z**. Ce qui est malgré les apparences, un progrès. Mais le principe (estimation par l'espérance), et le modèle, restent les mêmes.

(22) Quitte à revenir en une autre occasion sur la technique de Lyon. (Ainsi que sur certaines tentatives, malheureuses, de Metcalf). Il pourra être utile aussi d'examiner le modèle de Brunetti, qui est un peu différent. Et enfin d'évoquer la pratique statisticienne dans d'autres domaines que celui des monnaies : l'estimation du nombre des espèces d'après un échantillon a, en effet, préocupé plusieurs chercheurs, qui s'ignorent mutuellement.

Références

1) I.D. BROWN. Some notes on the coinage of Elizabeth I..., (The Bristish Numismatic Journal 1955-57, XXVIII, mcmlviij, p. 580).

2) J.B. GIARD, Les monnaies du premier consulat d'Octave, (Revue Numismatique, VIe série, tome XIII, année 1971, pp. 97-99, cf. p. 167, ibid.).

3) P. GRIERSON, The volume of Anglo-Saxon coinage, (The Economic History Review, second series, XX, nos 1-2-3, 1967, p. 153).

4) C.S. LYON, The estimation of the number of dies employed in a coinage, (The Numismatic Circular, Sept. 1965, vol. LXXIII, no 9).

5) D.M. METCALF, Offa's pence reconsidered, (Cunobelin, the yearbook of the British Association of Numismatic Society, 1963, no 9, p. 44).