



Une étude statistique des séries de deux ou plusieurs types d'éléments The statistical study of series of two or many types of elements: a digest

Louis Laurencelle ^a

^aUniversité du Québec à Trois-Rivières

Abstract ■ Series of outcomes, of conditions, of events regularly occur in our lives or they are encountered in our workplace, and sometimes they ought to be scrutinized. For example, we find at the factory that machine #2 produces every day from 3 to 5 defective artifacts out of 20, or that early in the morning at the hospital emergency clinic, out of 112 patients examined, 28 had severe symptoms of enteritis. The systematic mathematical study of series was, as we know, first addressed to games of chance, at the same time as was probability theory. For example, throwing a coin N times, what are the chances that we observe n Face, or that the series of N has r runs (or sequences with results of the same side), or that the longest run has L Face? Here we present a summary of the main results of the study of statistical series of N items falling into 2 or k (>2) categories, whether these items are provided from the outset or whether they emanate from a parameterized random process (binomial or multinomial). Formulas, examples and tables of critical values are provided.

Keywords ■ statistical series, runs, categorical variables, critical values, binomial variables, multinomial variables.

Acting Editor ■ Denis Cousineau (Université d'Ottawa)

Version révisée et corrigée d'une version antérieure parue en 2000 dans Lettres statistiques, vol. 11 (2000), p. 139-174.

louis.laurencelle@gmail.com

[10.20982/tqmp.16.4.p391](https://doi.org/10.20982/tqmp.16.4.p391)

Introduction

Lorsque des événements se produisent l'un après l'autre et peuvent être classés en deux ou plusieurs catégories ou types, leur enregistrement peut se faire sous la forme de la série de leurs types observés. C'est à l'étude de ces séries que nous nous attardons dans cet article. Prenons l'exemple de l'arrivée des 12 premiers amateurs de hockey sur glace à la porte de l'aréna, un soir donné. La liste ordonnée des personnes (P_i) pourrait être notée :

$\{P_1 P_2 P_3 P_4 P_5 P_6 P_7 P_8 P_9 P_{10} P_{11} P_{12}\}$

Si, maintenant, nous catégorisons chaque personne selon un aspect de son statut civil, soit Enfant (J, moins de 18 ans), Aîné(e) (A, 65 ans et plus), Homme (H) et Femme (F), nous obtenons la série suivante :

$\{H H J F H J H A A A F H\}$.

L'étude statistique des séries d'observations contenant deux à plusieurs types d'éléments en est encore à ses

débuts, malgré de belles percées en analyse combinatoire. Les chercheurs en éthologie quantitative, pour qui les séries d'observations catégorielles sont le pain quotidien, en sont largement encore à un stade descriptif et le traitement probabiliste de leurs chaînes de données reste décevant (LAURENCELLE, 1986).

Y a-t-il une structure statistiquement démontrable dans la série présentée ci-dessus ? À défaut d'une trousse d'analyse séquentielle adéquate, nous ne pourrions pas répondre à la question posée dans son sens général. Nous pouvons toutefois préciser notre question et demander par exemple s'il y a un type d'éléments qui domine en nombre un autre type, si les éléments de même type ont tendance à se grouper ou à se fuir, s'il y a un type d'éléments qui se groupent de façon extraordinaire. La réponse à la première question, sur le rapport entre les nombres d'éléments, est déjà connue et n'entre pas dans le cadre de ce texte. Nous tenterons plutôt d'étudier des questions pareilles aux deux



dernières, en identifiant à chaque fois le contexte d'observation et la question posée.

La section suivante précise le cadre et les notations de cet essai et présente les différentes quantités, ou variables, qui seront à l'étude. Nous passerons ensuite en revue, à tour de rôle, les séries et suites d'éléments binaires donnés, les séries d'éléments binaires libres, les séries d'éléments multiples donnés, enfin les séries d'éléments multiples libres. Rappelons que, par « suite », on désigne une portion de série ou séquence constituée d'éléments d'un même type; un élément d'un type donné précédé ou suivi d'éléments d'autres types constitue aussi une suite. À chaque occasion, référence sera faite aux sources documentaires consultées.

Notation, variables et cadres d'observation

Soit des éléments catégorisés en k types distincts, ou des boules identiques mais de k couleurs différentes. La série observée présente n éléments $X_i, i = 1$ à n , comportant n_1 éléments de type 1, n_2 de type 2, ..., n_k de type k , avec $n = n_1 + n_2 + \dots + n_k$. Pour illustration, chaque type d'élément est dénoté par une lettre minuscule, e.g. a, b , etc. Reprenons l'exemple donné plus haut, soit la série :

$$\{c \ c \ a \ d \ c \ a \ c \ b \ b \ b \ d \ c\}, \quad (1)$$

série qui comporte $n = 12$ éléments, dont $n_a = 2, n_b = 3, n_c = 5$ et $n_d = 2$.

Le nombre total de suites, dénoté S , correspond au nombre de séquences d'éléments de même type dans la série. En analysant la série (1), nous obtenons :

$$cc \ | \ a \ | \ d \ | \ c \ | \ a \ | \ c \ | \ bbb \ | \ d \ | \ c,$$

soit $S = 9$ suites, au total, i.e. 2, 1, 4 et 2 suites des types 1(= a), 2(= b), 3(= c) et 4(= d) respectivement. Dans l'illustration ci-dessus, une barre verticale sépare les éléments successifs de deux types différents : notons que le nombre de barres dénote aussi le nombre de changements de type et correspond au nombre S moins 1. La variable S prend des valeurs entières, depuis un minimum égal à k (ou $\sum n_j^0$) s'il y a k types différents représentés dans la série, jusqu'à un maximum égal à $\min\{n, 1 + 2[n - \max(n_i)]\}$.

Dans la série ci-dessus, on compte 7 suites de longueur 1, soit $L_1 = 7$, 1 de longueur 2 ($L_2 = 1$) et 1 de longueur 3 ($L_3 = 1$). La variable L_t satisfait les relations $L_t \geq 0$ et $\sum t \times L_t = n$; ici $1 \times 7 + 2 \times 1 + 3 \times 1 + 4 \times 0 + \dots = 12$.

Il est intéressant de considérer aussi les suites enchaînées, ou chaînes, de différentes longueurs. Nous appelons chaîne de longueur t une suite d'au moins t éléments de même type, que celle-ci soit enchâssée ou non dans une suite plus longue. La série (1) présente 1 chaîne de longueur 3 (« bbb »), ou $C_3 = 1$; 3 chaînes de longueur

2 ($C_2 = 3$), soit la chaîne propre « cc » et les deux sous-chaînes « bb. » et « .bb »; enfin, n chaînes de longueur 1 ($C_1 = n$).

La relation entre le nombre de chaînes d'une longueur donnée et le nombre de suites est facile à établir. Notons que chaque suite de longueur t donne une chaîne de même longueur; que chaque suite de longueur $t + 1$ donne 2 chaînes de longueur t (en utilisant les éléments 1 à t , puis 2 à $t + 1$); que chaque suite de longueur $t + 2$ donne 3 chaînes de longueur t , ainsi de suite. Il appert ainsi que :

$$C_t = \sum_{u \geq 1} u \times L_t - 1 + u. \quad (2)$$

Ainsi, dans notre exemple, nous obtenons $C_3 = 1 \times L_3 = 1$ et $C_2 = 1 \times L_2 + 2 \times L_3 = 1 \times 1 + 2 \times 1 = 3$. Notons l'ensemble des relations (2) en tableau, comme :

$$\begin{aligned} C_1 &= L_1 + 2 \times L_2 + 3 \times L_3 + 4 \times L_4 + 5 \times L_5 + \dots = n; \\ C_2 &= L_2 + 2 \times L_3 + 3 \times L_4 + 4 \times L_5 + \dots; \\ C_3 &= L_3 + 2 \times L_4 + 3 \times L_5 + \dots; \\ &\text{etc.} \end{aligned} \quad (1)$$

Nous remarquons que chaque C_t contient L_t , plus un excédent si $t > 1$. La contre-relation :

$$L_t = C_t - 2 \times C_{t+1} + C_{t+2} \quad (3)$$

restitue proprement le nombre de suites de longueur t à partir des nombres de chaînes. Pour notre exemple, avec la longueur de suite 1, nous obtenons $L_1 = C_1 - 2 \times C_2 + C_3 = 12 - 2 \times 3 + 1 = 7$.

La longueur maximale de suite observée ici est 3, ou $L_{max} = 3$. Notons que L_{max} prend des valeurs entières, depuis un minimum de 1 jusqu'à un maximum de $\max(n_j)$.

Finalement, nous pouvons nous intéresser aux répétitions, R , soit le nombre total d'apparitions d'un élément succédant à un autre de même type. Dans la série (1), nous dénombrons $R = 3$ répétitions, réparties en 0, 2, 1 et 0 répétitions selon le type. Une suite de longueur t contient $t - 1$ répétitions, d'où la correspondance :

$$R = \sum_{t \geq 1} (t - 1) \times L_t. \quad (4)$$

La relation (4) peut s'écrire $\sum (t \times L_t) - \sum 1$ ou, équivalamment, $n - S$, puisque $\sum_{t \geq 1} 1$ désigne le nombre total de longueurs de suites, i.e. le nombre de suites. Remarquant de plus que chaque chaîne de longueur 2 contient 1 répétition, nous obtenons l'égalité importante :

$$R = n - S = C_2. \quad (5)$$

Une autre quantité, les répétés (R'), ou « éléments répétés », désigne le nombre de fois que les types sont



répétés, i.e. le nombre d'occurrences d'éléments au-delà de leur première apparition. Dans une série donnée affichant k types d'éléments, nous avons évidemment $R' = n - k$, ou $12 - 4 = 8$ pour la série (1) en exemple. En général, nous observons $R' = n - \sum n_j^0$.

Deux cadres d'observation pour nos séries sont possibles. Premièrement, nous pouvons considérer la série $\{ X_i, i = 1 \text{ à } n \}$, comme étant donnée avec sa liste d'éléments et étudier l'une ou l'autre des variables de série sous l'univers des $n!$ permutations possibles de ses n éléments : c'est une série d'éléments donnés, que ceux-ci soient de 2 types (binaires) ou de $k \geq 2$ types (multiples). Deuxièmement, nous pouvons étudier un processus générateur d'éléments, chaque élément de type j ayant une probabilité constante π_j d'apparaître, indépendamment des éléments voisins. Ce processus produit à volonté des séries d'éléments, que nous appelons séries d'éléments libres. Ce processus peut être binomial, avec 2 types d'éléments, selon le vecteur générateur :

$$\begin{pmatrix} a & b \\ \pi & 1 - \pi \end{pmatrix} \tag{6}$$

et, pour des séries de n éléments, avec les probabilités d'apparition des éléments de chaque type :

$$B(n_1, n_2 | n, \pi, 1 - \pi) = \binom{n_1 + n_2}{n_1} \pi^{n_1} (1 - \pi)^{n_2}, \tag{7}$$

où $\binom{n}{x}$ dénote le nombre de combinaisons de n objets pris x à la fois, ou $n!/[x!(n-x)!]$. Le processus général est multinomial, avec $k \geq 2$ types d'éléments, selon le vecteur générateur :

$$\begin{pmatrix} a & b \dots & k \\ \pi_a & \pi_b \dots & \pi_k \end{pmatrix} \tag{8}$$

et, pour des séries de n éléments, avec les probabilités d'apparitions :

$$M(n_1, n_2, \dots, n_k | n, \pi_1, \pi_2, \dots, \pi_k) = \binom{n}{(n_j)} \pi_1^{n_1} \times \pi_2^{n_2} \times \dots \times \pi_k^{n_k}, \tag{9}$$

où $\binom{n}{(n_j)}$ est le coefficient multinomial, égal à $n!/[n_1! \times n_2! \times \dots \times n_k!]$.

L'article de A. M. Mood paru en 1940, The distribution theory of runs, est une source fondamentale sur les mathématiques des suites. Les théorèmes et résultats les plus nombreux, sur les séries et les suites, touchent surtout deux parmi les quatre catégories de séries ci-dessus mentionnées, soit les séries d'éléments binaires donnés et les

séries d'éléments multiples libres. Bien que plusieurs questions demeurent sans réponse au sujet de quelques-unes des variables d'étude énumérées, les données disponibles, publiées ou non, sont néanmoins abondantes. C'est pourquoi nous nous contenterons de présenter ici les résultats qui nous paraissent les plus intéressants ou les plus utiles, sans en reproduire la démonstration.

Les séries d'éléments binaires donnés

Nous observons une série d'éléments de deux types, par exemple une file d'hommes et de femmes à l'arrêt d'autobus, le signe positif ou négatif de l'écart (ou résidu) d'une valeur mesurée par rapport à la fonction de régression qui lui sert de modèle, le côté apparent, face ou pile, d'un jeton lancé n fois. Nous comptons n_1 éléments de type a , n_2 de type b , avec $n = n_1 + n_2$. Les n éléments donnent lieu à $n!$ permutations; parmi celles-ci, les éléments de même type peuvent être interchangés sans modifier les propriétés de la série ni la valeur des variables mesurées. Il reste donc $\binom{n}{n_1}$ arrangements mutuellement discriminables, tous équiprobables quant au nombre d'éléments de chaque type, mais qu'en est-il de la probabilité des suites, de leurs longueurs, des répétitions d'éléments ?

Le nombre total de suites, S

Le nombre total de suites, S , partant d'un minimum de 2, peut aller jusqu'à n , si $n_1 = n_2$, ou jusqu'à $1 + 2 \times \min(n_1, n_2)$ si $n_1 \neq n_2$. MOOD (1940), BRADLEY (1968) et d'autres démontrent que la distribution de probabilités de S obéit à la paire de formules suivante, selon que S est pair ou impair :

$$p(S = 2u) = \frac{2 \binom{n_1}{u-1} \binom{n_2}{u-1}}{\binom{n}{n_1}}; \tag{10}$$
$$p(S = 2u + 1) = \frac{\binom{n_1-1}{u-1} \binom{n_2-1}{u} + \binom{n_1-1}{u} \binom{n_2-1}{u-1}}{\binom{n}{n_1}}.$$

Les quatre premiers moments sont donnés dans Laurencelle (1995; voir aussi Laurencelle et Dupuis, 2000) et sont :

$$\mu(S) = 1 + \frac{2n_1n_2}{n}; \tag{11a}$$

$$\sigma^2(S) = \frac{2n_1n_2(2n_1n_2 - n)}{n^2(n-1)} = \frac{(\mu-1)(\mu-2)}{n-1}; \tag{11b}$$

$$\gamma_1(S) = \sqrt{\frac{n-1}{2(n-2)^2}} \times \frac{16n_1^2n_2^2 - 4n_1n_2n(n+3) + 3n^3}{\sqrt{n_1n_2(2n_1n_2 - n)^3}}; \tag{11c}$$



$$\gamma_2 = \frac{(n-1)[n^4(8n_1n_2-13) - 7n^5 + 2n^3(52n_1n_2+3) - 120n^2n_1^2n_2^2 + 24n \times n_1^2n_2^2(n_1n_2-6) + 144n_1^3n_2^3]}{2n_1n_2(2n_1n_2-n)^2(n-2)(n-3)} - 3 \quad (11d)$$

WEISS (1988) propose une dérivation très simple de l'espérance (11a).

Des tables de valeurs critiques de S , pour $2 \leq n_1, n_2 \leq 20$ et $21 \leq n_1 = n_2 \leq 100$, se trouvent un peu partout, par exemple dans OWEN (1962); la plupart proviennent de SWED et EISENHART (1943). Malgré l'ultime normalité de S pour un rapport n_1/n_2 constant et n croissant, la distribution est négativement asymétrique, l'indice γ_1 ayant pour facteur $-(n_1 - n_2)^2$, de sorte que l'approximation normale simple reste boiteuse. LAURENCELLE et DUPUIS (2000) proposent une meilleure approximation, basée sur une série de Edgeworth à trois termes (voir plus bas), de même qu'un jeu complet de valeurs critiques correspondant aux percentiles 0,005, 0,01, 0,025, 0,05, 0,95, 0,975, 0,99 et 0,995, ce pour $2 \leq n_1, n_2 \leq 50$.

Une autre statistique des séries, soit le nombre de transitions (ou variations) d'un type à l'autre d'élément, est liée naturellement au nombre S de suites et égale à $S - 1$.

Le nombre de suites pour l'élément « a », S_1

On peut s'intéresser aussi à l'un seul des deux éléments de la série binaire, disons l'élément a . S'il y a n_1 éléments de ce type, n_2 de l'autre et $n = n_1 + n_2$, le nombre de suites S_1 des éléments a varie entre $\min(1, n_1)$ et $\min(n_1, \lfloor (n + 1)/2 \rfloor)$. La distribution de probabilités de S_1 est :

$$p(S_1) = \frac{\binom{n_1-1}{S_1-1} \binom{n_2+1}{S_1}}{\binom{n}{n_1}} \quad (12)$$

avec, pour moments :

$$\mu(S_1) = \frac{n_1(n_2+1)}{n}; \quad (13a)$$

$$\sigma^2(S_1) = \frac{n_1n_2(n_1-1)(n_2+1)}{n^2(n-1)} = \frac{\mu(\mu-1)}{n-1}; \quad (13b)$$

$$\gamma_1 = \frac{4n_1(n_2+1) - n(n+2)}{\sqrt{\frac{n_1n_2(n_1-1)(n_2+1)(n-2)^2}{n-1}}}; \quad (13c)$$

$$\gamma_2 = \frac{6n_1^4(5n-6) - 12n_1^3(n+1)(5n-6) + 6n_1^2(7n^3+7n^2-13n-6) + 6n_1n(n+1)(2n^2+3n-6) - n^2(n-1)(n+1)(n+6)}{n_1n_2(n_1-1)(n_2+1)(n-2)(n-3)} \quad (13d)$$

La longueur maximale de suite, L_{max}

(BRADLEY, 1968, p. 255 et suiv.) développe ingénieusement une expression fournissant la probabilité

que la série comporte au moins une suite d'au moins L ($= L_{max}$) éléments d'un type ou l'autre. L'expression globale est :

$$Pr\{L_{max} \geq L\} = Pr\{L_{max_1} \geq L\} + Pr\{L_{max_2} \geq L\} - Pr\{L_{max_1} \geq L \text{ et } L_{max_2} \geq L\}, \quad (14)$$

où L_{max_j} réfère à une suite d'éléments de type j ($j = 1$ ou 2). Les ingrédients sont :

$$Pr\{L_{max_1} \geq L\} = \binom{n}{n_1}^{-1} \sum_{i=1}^{n_1/L} (-1)^{i+1} \binom{n_2+1}{i} \binom{n-i \times L}{n_2}, \quad (15a)$$

$$Pr\{L_{max_2} \geq L\} = \binom{n}{n_1}^{-1} \sum_{i=1}^{n_2/L} (-1)^{i+1} \binom{n_1+1}{i} \binom{n-i \times L}{n_1}, \quad (15b)$$



$$Pr\{L_{max_1} \geq L \text{ et } L_{max_2} \geq L\} = \binom{n}{n_1}^{-1} \sum_{s_1=1}^{n_1-L+1} \{A(s_1) [B(s_1) + 2C(s_1) + D(s_1)]\}, \quad (15c)$$

où :

$$A(s_1) = \sum_{i=1}^{(n_1-s_1)/(L-1)} (-1)^{i+1} \binom{s_1}{i} \binom{n_1-1-i \times (L-1)}{s_1-1};$$

$$B(s_1) = \sum_{i=1}^{(n_2-s_1+1)/(L-1)} (-1)^{i+1} \binom{s_1-1}{i} \binom{n_2-1-i \times (L-1)}{s_1-2};$$

$$C(s_1) = \sum_{i=1}^{(n_2-s_1)/(L-1)} (-1)^{i+1} \binom{s_1}{i} \binom{n_2-1-i \times (L-1)}{s_1-1};$$

$$D(s_1) = \sum_{i=1}^{(n_2-s_1-1)/(L-1)} (-1)^{i+1} \binom{s_1+1}{i} \binom{n_2-1-i \times (L-1)}{s_1}.$$

À partir de (14), la probabilité d'obtenir la longueur maximale L est évidemment $p_{L_{max}}(L) = Pr\{L_{max} \geq L\} - Pr\{L_{max} \geq L + 1\}$. Nous n'avons pas trouvé d'expression pour les moments de L_{max} . L'étude des valeurs calculées de l'espérance, cependant, montre que, 1) pour n_1 constant et $n_2 > n_1$, l'espérance de L_{max} croît linéairement avec n_2 , et que 2) pour $n_1 = n_2$, l'espérance augmente comme le logarithme de n_1 (ou n_2), en fait $\mu(L_{max}) \approx 0,039 + 1,7 \times \log_e(n_1 + 0,7)$.

Il n'y a pas, à notre connaissance, de tables assez complètes de valeurs critiques pour L_{max} . C'est pourquoi nous avons préparé le tableau 1 à la fin de cet article, qui permettent de décider si, aux différents seuils de probabilité 0,05, 0,01 et 0,001, la longueur de suite maximale L_{max} dans une série est significative d'un groupement trop important pour être attribué au hasard, ce pour les valeurs $1 \leq n_1 \leq n_2 \leq 100^+$ et $n_2 - n_1 \leq 10$.

Autres aspects de la série

Les séries d'éléments binaires donnés, qui sont un cas particulier des séries d'éléments multiples donnés, comportent des aspects autres que le nombre ou la longueur maximale des suites. Nous étudierons ces aspects avec le cas général d'éléments multiples.

Un exemple

L'exemple qui suit est repris de LAURENCELLE (1995). Supposons une machine qui produit des items en série, au rythme d'un item par unité de temps. Quelques-uns des items produits ont un défaut, d'origine inconnue. L'ingénieur en charge prend note d'une série de $n = 46$ items, laquelle contient 8 items défectueux. Ces défauts se produisent-elles au hasard ou ont-elles tendance à apparaître par groupes? L'examen de la série des 46 items permet de dénombrer ($S = 9$) suites, au total : ces suites sont-elles trop peu nombreuses pour que l'apparition des défauts soit imputée au hasard? Afin de répondre à la question au seuil de signification de 0,01, il suffit d'établir $P(9) = p(S \leq 9)$, exactement ou en approximation, puis vérifier si $P(9) \leq 0,01$.

Les formules (11a) à (11d) des moments de S , pour $n_1=8$ items défectueux et $n_2 = 38$ corrects, permettent d'établir l'espérance (μ) 14,2174, la variance (σ^2) 3,5885 et l'écart-type (σ) 1,8943, l'indice d'asymétrie (γ_1) -0,4503 et l'indice d'aplatissement (γ_2) -0,0955. L'approximation normale, avec correction de continuité, a pour formule $z = (S \pm \frac{1}{2} - \mu)/\sigma$ et $P(S) \approx \Phi(z)$, $\Phi(\cdot)$ étant la fonction de répartition de la loi normale standard. Nous calculons $z = (9 + \frac{1}{2} - 14,2174)/1,8943 \approx -2,490$. Puisque $\Phi(-2,490) \approx 0,00639$, on déclarerait que le nombre de suites, $S = 9$, est significativement bas et que les anomalies



de production tendent à apparaître par groupes. L'autre approximation, que nous recommandons lorsque $n_1 \neq n_2$, se base sur le même écart-réduit z , puis $P(S) \approx \Phi(z) - \phi(z) \left[\frac{\gamma_1}{6} (z^2 - 1) + \frac{\gamma_2}{24} (z^3 - 3z) + \frac{\gamma_3}{72} (z^5 - 10z^3 + 15z) \right]$, $\phi(\cdot)$ étant la densité normale standard. Utilisant encore $z = -2,490$, nous obtenons ici $P(9) = \Phi(-2,490) - \phi(-2,490) \times [\gamma_1(0,8667) + \gamma_2(-0,3320) + \gamma_{12}(0,2960)] \approx 0,0118$, un résultat non significatif au seuil choisi de 0,01. Le calcul exact de $P(9)$, à l'aide de la distribution (10), donne une probabilité cumulative de 0,01283, donc un résultat non significatif. Enfin, la table publiée dans LAURENCELLE et DUPUIS (2000, p. 161-169) donne, pour $\alpha = 0,01$, $n_1 = 8$ et $n_2 = 38$, la valeur critique 8 : notre résultat, $S = 9$, n'est décidément pas significatif : au seuil de 0,01, on ne peut pas affirmer que les défauts constatés dans la production d'items en série tendent à être séquentiellement groupés.

Les séries d'éléments binaires libres

Nous sommes ici face à un processus générateur produisant des éléments de deux types, disons les éléments a et b , avec probabilités respectives π et $1 - \pi$. Chaque production est un essai de Bernoulli, et une série de productions constitue une série de Bernoulli, en admettant que les probabilités sont constantes et que chaque production est indépendante des autres. Ainsi, dans une série anticipée de n essais de Bernoulli, les nombres n_1 et n_2 d'éléments des deux types sont des variables aléatoires, sujettes à $n_1 + n_2 = n$, en plus de l'ordre des éléments de la série et des variables qui en découlent. Les nombres binomiaux n_1, n_2 obéissent à la loi binomiale (7), disons $p(n_1, n_2 | n, \pi)$.

Le nombre total de suites, S

La distribution de S , le nombre total de suites, dépend bien entendu des nombres n_1 et n_2 d'éléments présents. Pour n_1 et n_2 donnés, les expressions (10) fournissent la probabilité conditionnelle, $p(S | n_1, n_2)$ d'obtenir S suites dans la série. La probabilité cherchée ici est celle d'observer S suites dans une série quelconque de n essais de Bernoulli, à probabilités π et $1 - \pi$, soit $p(S | n, \pi)$. On obtient cette probabilité en convoluant la distribution de S avec celle de n_1, n_2 , c'est-à-dire :

$$p(S | n, \pi) = \sum_{n_1=0}^n p(S | n_1, n - n_1) \times p(n_1, n - n_1 | n, \pi). \tag{16}$$

Ce calcul, assez laborieux, a permis d'établir les tables 2 et 3 de valeurs critiques de S à différents seuils de signification, ce pour $\pi = 1/2, 1/3, 1/4, 1/5, 1/6$ et $1/10$ et $n \leq 50$. Pour $S = 1$, nous avons simplement $p(S = 1 | n, \pi) = \pi^n + (1 - \pi)^n$. Enfin, le cas particulier d'éléments équiprobables, i.e. pour

lesquels $\pi = 1 - \pi = \frac{1}{2}$, est repris avec le cas général d'éléments multiples libres équiprobables.

Une autre statistique des séries, soit le nombre de transitions (ou variations) d'un type à l'autre d'élément, est naturellement liée au nombre S de suites et égal à $S - 1$.

MOOD (1940) donne les ingrédients permettant d'établir les moments de S , dans le cas des séries binaires libres. Les deux premiers moments sont :

$$\mu(S | n, \pi) = 1 + 2\pi(1 - \pi)(n - 1); \tag{17a}$$

$$\sigma^2(S | n, \pi) = \pi(1 - \pi)[4(n - 1)(1 - 3\pi(1 - \pi)) - 2(1 - 2\pi)^2]. \tag{17b}$$

Considérant l'espérance (11a) de S applicable à des valeurs fixes de n_1 et $n_2 (= n - n_1)$, l'espérance (17a) peut s'obtenir par $E_{B(n, \pi)}\{E(S | n_1, n - n_1)\} = E_{B(n, \pi)}\{1 + 2n_1(n - n_1)/n\}$, en utilisant $E_{B(n, \pi)}(n_1) = n \times \pi$ et $E_{B(n, \pi)}(n_1^2) = \sigma^2(n_1) + E^2(n_1) = n \times \pi(1 - \pi) + (n \times \pi)^2$ [$E_{B(n, \pi)}$ dénote ici l'espérance selon la loi binomiale].

Le nombre de suites pour l'élément a, S_1

Considérant à présent seules les suites de l'élément a , la distribution de probabilités $p(S_1 | n, \pi)$ du nombre de suites (S_1) pour cet élément s'obtient comme ci-dessus, par une convolution de (12) avec (7). Ceci donne lieu aussi à des valeurs critiques, qu'on retrouve aux tables 4 et 5 : ces valeurs concernent les suites (S_1) de l'élément plus rare (a) dans une série de Bernoulli, celui pour lequel $p(a) = \pi \leq 1 - \pi$, et peuvent intéresser celui qui, par exemple, s'inquiète de la plus ou moins grande périodicité des apparitions de a . Noter qu'il est ici possible d'observer $S_1 = 0$ suite de l'élément a , ce qui se produit lorsque seuls des éléments b sont présents, selon la probabilité $(1 - \pi)^n$. Quant à la probabilité d'observer 1 suite d'éléments a ou moins, soit $p(S_1 \leq 1)$, sa valeur pour $\pi < \frac{1}{2}$ est égale à $\pi^n + (1 - \pi)^n + 2[\pi \times (1 - \pi)^n - \pi^n(1 - \pi)] / (1 - 2\pi) + \pi \times (1 - \pi)^n \{[n - 2 - \pi \times (2n - 3)] + \pi^n(1 - \pi)^2\} / (1 - 2\pi)^2$ et, pour $\pi = \frac{1}{2}$, $n \times (n + 1) / 2^{n+1}$.

Les deux premiers moments de S_1 dans ce contexte d'éléments binaires libres sont :

$$\mu(S_1 | n, \pi) = n \times \pi(1 - \pi) + \pi^2; \tag{18a}$$

$$\sigma^2(S_1 | n, \pi) = n \times \pi(1 - 4\pi + 6\pi^2 - 3\pi^3) + \pi^2(3 - 8\pi + 5\pi^2). \tag{18b}$$

Autres aspects de la série

La distribution de la longueur maximale de suite (L_{max}) pour des éléments libres peut être obtenue par la même méthode de convolution appliquée ci-dessus au nombre de suites. En fait, l'expression (14), qui donne une probabilité cumulative applicable aux séries comptant n_1 et n_2 éléments de chaque type, doit être convoluée avec la probabilité binomiale (7), puis décumulée pour fournir



enfin $p_{L_{max}}(L|n, \pi)$. Il en est de même pour L_{max_1} , la longueur maximale de suite de l'élément a , en utilisant cette fois (15a).

D'autres mesures intéressantes sont possibles avec les séries d'éléments binaires libres, comme les répétitions, les répétats, etc. Nous les considérons plus loin, avec le cas général d'éléments multiples équiprobables.

Un exemple

Nous lançons un dé en l'air 50 fois, et il retombe 12 fois à la face «1»; la face « 1 » apparaît en fait dans 5 suites, soit deux suites d'un seul élément ($L_1 = 2$), une suite de 2 éléments « 1 » consécutifs ($L_2 = 1$), une de 3 ($L_3 = 1$) et une de 5 ($L_5 = 1$). Y a-t-il quelque chose de remarquable dans cette expérience ?

En admettant que nous ayons un dé honnête à six faces, la probabilité d'occurrence de chacune étant 1/6, les résultats observés constituent une série de Bernoulli dans laquelle la face «1» constitue l'événement d'intérêt et où les autres faces du dé sont confondues en «l'autre élément» et présentent conjointement une probabilité de 5/6. Dans ce

cas, la probabilité d'obtenir 12 fois la face «1» est, d'après (7), de $B(12, 38|50, 1/6, 5/6) = 0,05464$; on peut aussi calculer $Pr(x \leq 12) \approx 0,93733$ et $Pr(x \geq 12) \approx 0,11731$. Donc, exception faite d'un léger excédent du nombre de «1» par rapport au nombre attendu, $n \times \pi \approx 8,333$, le nombre de faces «1» obtenu n'est pas vraiment exceptionnel.

Les 12 faces « 1 » dans notre expérience sont apparues en $S_1 = 5$ suites. Aux tableaux 4 et 5, un tel résultat n'apparaît remarquablement ni bas ni élevé : la probabilité d'obtenir 5 suites ou moins, en contexte de Bernoulli, est ici de 0,23642. Par contre, selon la table, l'obtention de $S_1 = 3$ suites ou moins permettrait de juger que les 50 lancers du dé produisent trop peu de suites de l'élément « 1 », cet élément ayant tendance à se grouper.

Quant aux longueurs des suites, la longueur maximale (L_{max}) observée est de 5 : est-ce remarquable ? Pour répondre à cette question, il faudrait trouver $Pr(L_{max_1} \geq 5)$ en contexte de Bernoulli. L'utilisation de (15a) en convolution avec (7) fait l'affaire, à savoir :

$$Pr(L_{max_1} \geq L) = \sum_{n_1=L}^n \pi^{n_1} (1-\pi)^{n-n_1} \left(\sum_{i=1}^{n/L} (-1)^{i+1} \binom{n-n_i+1}{i} \binom{n-i \times L}{n-n_1} \right).$$

Dans notre cas, nous posons $n = 50$, $\pi = 1/6$, $L = 5$ et nous calculons alors 0,00494; cette infime probabilité révèle le caractère exceptionnel d'une suite de 5 faces consécutives dans 50 lancers d'un dé.

Nous aurions pu aussi considérer les résultats observés comme éléments donnés, sans référence au processus générateur et aux probabilités (e.g. $\pi = 1/6$) qui s'y attachent. Dans ce cas, la probabilité que nos 12 faces « 1 », observées parmi 38 autres faces du dé, forment 5 suites est fournie par (12), soit ici $p(S_1 = 5) \approx 0,00157$. En fait, $Pr(S_1 \leq 5) \approx 0,00168$, l'arrangement observé manifestant un nombre remarquablement petit de suites, au seuil de signification de 1%. La comparaison avec la distribution de S_1 sous le contexte d'un processus de Bernoulli est ici surprenante. Quant à L_{max} , le calcul direct de (15a) fournit $Pr(L_{max_1} \geq 5) \approx 0,08314$, une valeur qui n'atteint même pas le seuil de 5% : dans une série comportant 12 faces «1» et 38 autres faces, il n'est pas vraiment surprenant d'obtenir ici une suite de 5 faces «1» consécutives.

Les séries d'éléments multiples donnés

L'analyse des séries d'éléments multiples, i.e. des séries de n éléments comportant $n_j \geq 0$ éléments de chaque type j , $1 \leq j \leq k$, où $n = \sum n_j$, est évidemment plus complexe. Avec $n_j > 0$ éléments donnés de chaque type, l'ensemble considéré englobe $n!$ arrangements d'éléments, parmi lesquels seuls $M_k(n; n_j) = \binom{n}{n_j}$ arrangements sont distincts; pour des éléments de deux types, ou éléments binaires, nous prenons $k = 2$ et $M_2(n; n_j) = \binom{n}{n_1}$. En fin de text, à l'encadré 1, le lecteur trouvera un programme en Pascal-Delphi qui énumère les $M_k(n; n_j)$ arrangements distincts et permet d'en analyser en détail le contenu. Le nombre multinomial $M_k(n; n_j)$ s'élève très vite avec k et le poids du vecteur (n_1, n_2, \dots, n_k) , de sorte que l'approche énumérative devient irréaliste, sauf pour de petites configurations spécifiques.

Le nombre total de suites, S

Soit $T(S|n_1, n_2, \dots, n_k)$ ou $T_k(S|n_j)$, le nombre total d'arrangements distincts des $n_j > 0$ éléments de différents types dans lesquels on retrouve S suites d'éléments sem-



blables, et soit la distribution de probabilités $p(S) = p_k(S|n_j) = T_k(S|n_j)/M_k(n; n_j)$. DAVID et BARTON (1962) présentent une méthode récursive permettant de calculer $T_k(S|n_j)$. SHAUGHNESSY (1981) présente aussi une

méthode et donne, en fait, trois fonctions récursives. La première récurrence présentée (op. cit., p. 734) est homogène et a la forme :

$$T(v+r|n_1, n_2, \dots, n_v) = \left[\sum_{t=1}^{r+1} [T(v+r-t|n_1, n_2, \dots, n_v-1) \times \sum_{j=0}^{n-n_v-v+1} \binom{n-n_v-v-r+t}{j} \binom{n_v-1}{t-1-j} \binom{v+r-t+1}{t-2j}] \right] \quad (19)$$

D'abord, pour trouver $T_k(S|n_j)$ depuis le niveau $v = k$, c'est-à-dire avec l'invocation $T(S|n_1, n_2, \dots, n_v)$ et $n = \sum_{j=1}^v n_j$, la descente récursive opère en réduisant v à $k-1$, puis $k-2$ etc., chaque fois avec l'invocation $T(S|n_1, n_2, \dots, n_v)$ et $n = \sum_{j=1}^v n_j$. Lorsqu'enfin $v = 1$, $T(S|n_1)$, qui dénote le nombre d'arrangements de n_1 éléments produisant S suites, vaut 1 pour $S = 1$ et vaut 0 dans tous les autres cas. L'encadré 2, en fin de texte présente un programme en Pascal-Delphi qui concrétise ces opérations. Il est intéressant de noter que $T_k(k|n_j) = k!$.

Le nombre de suites S varie d'un minimum égal à k à un maximum égal à $\min[n, 1 + 2(n - n_j \max)]$. MOOD (1940) et DAVID et BARTON (1962) donnent les ingrédients permettant d'obtenir les moments de S . Les deux premiers sont :

$$\mu(S) = n + 1 - \frac{\sum n_j^2}{n}; \quad (20a)$$

$$\sigma^2(S) = \frac{(\sum n_j^2)^2 + n(n+1)\sum n_j^2 - n(n^2 + 2\sum n_j^3)}{n^2(n-1)} \quad (20b)$$

Si les n éléments de k types sont en nombres égaux, i.e. si $n_j = n/k$ pour tout j , l'expression des moments se simplifie :

$$\mu(S) = n + 1 - n/k; \quad (21a)$$

$$\sigma^2(S) = \frac{n(n-k)(k-1)}{(n-1)k^2} = \frac{(\mu-1)(n-\mu)}{n-1}; \quad (21b)$$

$$\gamma_1(S) = -\sqrt{\frac{n(n-1)}{(n-k)(k-1)}} \times \frac{k-2}{n-2}; \quad (21c)$$

$$\gamma_2(S) = \frac{(n^3 - n + 6) \times k^2 - 6n \times k \times (n^2 - n + 2) + 6n^3}{n(n-2)(n-3)(n-k)(k-1)} \quad (21d)$$

La confection de tables de valeurs critiques se heurte avant tout à l'embarrassante multiplicité des configurations d'éléments. Ainsi, avec $n = 10$ et $k = 3$, il y a $P_a(n, k) = P_a(10, 3) = 8$ partitions¹ (ou configurations) possibles des nombres d'éléments, soit les partitions $n_1 n_2 n_3 = 811, 721, 631, 622, 541, 532, 442$ et 433 : chaque partition engendre sa propre distribution de probabilités. Assignant une modeste moyenne de 5 éléments par type et appliquant $k = 3, 4, \text{etc.}$, nous trouvons $P_a(15, 3) = 19$, $P_a(20, 4) = 64$, $P_a(25, 5) = 192$, voire $P_a(50, 10) = 16928$ partitions. En pratique, à notre avis, la meilleure solution consiste à faire une évaluation ad hoc de la distribution $p(S)$ pour chaque cas particulier, en utilisant le programme de calcul présenté en fin de texte dans l'encadré 1.

Le nombre de suites pour l'élément a , S_1

Le nombre de suites des éléments d'un seul type, disons les éléments a , dans une série à éléments multiples, obéit à la même distribution que dans une série binaire. Soit n_1 , le nombre d'éléments du type considéré, et redéfinissant $n_2 = n - n_1$, i.e. n_2 étant le nombre total d'éléments des autres types, la distribution (12) et les moments (13a)-(13d) décrivent le nombre de suites de l'élément a , la variable S_1 .

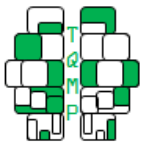
Les répétitions (R) et les chaînes de 2 éléments (C_2)

Répétitions (R) et chaînes de longueur 2 (C_2) sont des variables complémentaires du nombre total de suites (S), par la relation (5).² Leurs distributions sont semblables et en miroir de celle de S . Notons pour mémoire que leurs deux premiers moments sont :

$$\mu(R) = \mu(C_2) = \frac{\sum n_j^2}{n} - 1; \quad (22a)$$

1. Une partition est une décomposition additive d'un nombre naturel en un ou plusieurs autres naturels, p. ex. $P_a(5) = 5, 41, 32, 311, 221, 2111, 11111$, ces partitions contenant 2 partitions de 3 éléments, $P_a(5, 3) = 311, 221$. Cette fonction peut être calculée par récursion, selon $P_a(n, k) = \sum_{i=1}^k P_a(n-i, i)$, avec $P_a(n, 1) = 1$ et $P_a(n, n+r) = P_a(n, n)$, $r \geq 0$.

2. Utilisant la relation (2), nous avons en effet $C_2 = L_2 + 2L_3 + 3L_4 + \dots$. Or, $L_1 + 2L_2 + 3L_3 + \dots = n$, d'où $n - C_2 = (L_1 + 2L_2 + 3L_3 + \dots) - (L_2 + 2L_3 + 3L_4 + \dots) = L_1 + L_2 + L_3 + \dots = S$, CQFD.



$$\sigma^2(R) = \sigma^2(C_2) = \frac{(\sum n_j^2)^2 + n(n+1)\sum n_j^2 - n(n^2 + 2\sum n_j^3)}{n^2(n-1)} \tag{22b}$$

les mêmes simplifications, données en (21a)-(21d) pour le nombre total de suites, s'appliquent ici. Rappelons toutefois que, les variables R et C_2 étant complémentaires à S , le signe de leur moment d'asymétrie γ_1 devient positif.

La longueur maximale de suite, L_{max}

Les auteurs (MOOD, 1940; DAVID & BARTON, 1962) ne fournissent pas de méthode explicite permettant de

déterminer la distribution de probabilités de la longueur de suite maximale (L_{max}) ni ses moments (voir cependant David et Barton, op. cit., p. 232-238).

Posant une configuration d'éléments n_1, n_2, \dots, n_k dans laquelle $n_1 \geq n_2 \geq \dots \geq n_k$, la longueur maximale de suite varie d'un minimum égal à $\max[1, \lceil n_1/(n - n_1 + 1) \rceil]$ jusqu'à n_1 comme maximum ($\lceil x \rceil$ désigne le plus proche entier non inférieur à x). La distribution de probabilités, de forme $p(L_{max}) = C(L_{max}|n_j)/M_k(n; n_j)$, n'est pas connue et doit être déterminée par énumération.³ Toutefois, nous avons trouvé l'expression suivante pour la valeur la plus élevée de L_{max} , soit $L_{max} = n_1$:

$$C(n_1|n_j) = \sum_{i=1}^t (-1)^{i+1} \times i! \binom{t}{i} \binom{n - \sum_{j=1}^i n_j + i}{i} \frac{(n - \sum_{j=1}^i n_j)!}{\prod_{j=i+1}^k n_j!}, \tag{23}$$

expression dans laquelle t dénote le nombre de valeurs n_j égales à n_1 , i.e. tel que $n_1 = n_2 = \dots = n_t, t \geq 1$. De plus, admettant $n_1 = \max[n_j, j > 1] + d, d > 0$, on obtient facilement le nombre $C(n_1 - u|n_j)$, pour $u \leq \min(d - 1, n_1/2 + 1)$, en utilisant :

$$C(L_{max} \geq n_1 - u|n_j) = \frac{n - n_1 + 1}{u!} \times \frac{(n - n_1 + u)!}{\prod_{j=2}^k n_j!}, \tag{24}$$

puis en calculant $C(n_1 - u) = C(L_{max} \geq n_1 - u) - C(L_{max} \geq n_1 - u + 1)$. Nous n'avons pas poussé plus loin notre exploration.

La longueur maximale de suite pour l'élément a , L_{max1}

Comme pour le nombre de suites (S_1) de l'élément a , on peut rapporter l'étude de la longueur maximale de suite d'un élément quelconque, désigné a , en réunissant les $k - 1$ éléments restants, constituant ainsi une série binaire de n_1 éléments d'un type donné et $n - n_1 = n_2$ éléments complémentaires. La formule (15a) déjà donnée s'applique alors.

Autres aspects de la série

Il est théoriquement possible d'obtenir la distribution $p(L|n_j)$ de la longueur de suites, nommément la probabilité qu'une série prise au hasard parmi les $M_k(n|n_j)$ séries possibles présente une à plusieurs suites de longueur L . La littérature est muette sur cette question. Enfin, dans les séries à éléments donnés, le nombre de répétats (R') n'est

pas variable et il est toujours égal à $n - k$.

Un exemple

Dans une usine, on fabrique quatre pièces d'équipement majeures, disons a, b, c et d . Chaque pièce est fabriquée sur demande. Durant trois semaines, le contremaître note la série de pièces commandées, soit

a b a a a a a d a b
c a a b a b c a a d.

Il y a $n_j = 12, 4, 2$ et 2 commandes par type de pièces. Ces commandes arrivent-elles en désordre ou, sinon, y aurait-il avantage à spécialiser des sous-groupes de techniciens dans la fabrication d'un type de pièces particulier? La série comptabilisée comporte $S = 14$ suites, la longueur de suite maximale (L_{max}) étant 5.

Compte tenu des nombres (n_j) d'éléments observés, les moments du nombre de suites, par (18a)-(18b), sont $\mu(S) = 12,600$ et $\sigma^2(S) \approx 2,429$. Par la fonction (19) ou le programme donné en fin de texte à l'encadré 2, nous calculons $p(S = 14) \approx 0,18783, Pr(S \leq 14) \approx 0,89600$ et $Pr(S \geq 14) \approx 0,29183$: rien de remarquable là-dedans. Quant aux longueurs de suites, comme $n_1 > L_{max} \geq n_2/2 + 1$, nos formules (23) et (24) ne s'appliquent pas. Utilisant le programme d'énumération, à l'encadré 2, nous obtenons $C(L_{max} = 5) = 12307680$ et $Pr(L_{max} = 5) = C(L_{max} = 5)/M_4(20; 12, 4, 2, 2) = 12307680/52907400 \approx 0,23263$ et, de même, $Pr(L_{max} \geq 5) \approx 0,44689$. Le calepin de commandes ne laisse donc

3. Dans le programme d'énumération présenté en fin de texte dans l'encadré 1, il suffit de réécrire la procédure « Analyser » afin de repérer dans chaque série la suite maximale et d'enregistrer sa longueur dans le tableau « hist[...] » prévu à cet effet.



voir aucune tendance véritable au groupement, et le travail des techniciens d'usine gagne à rester polyvalent.

Les séries d'éléments multiples libres

Au lieu d'observer une série d'éléments déterminés dont seul l'ordre séquentiel est en jeu, nous nous plaçons devant un générateur multiple, un processus multinomial, capable de produire des séries de n éléments de k types selon le vecteur de probabilités (8). Nous tenterons d'estimer la probabilité de, ou prédire, certains aspects des séries ainsi engendrées. Dans le cas général, nous avons affaire à un processus à probabilités $\{\pi_1 \pi_2 \dots \pi_k\}$ quelconques, cas pour lequel peu de résultats mathématiques sont disponibles. Le cas spécial d'équiprobabilité, $\{1/k 1/k \dots 1/k\}$, plus documenté, est traité ensuite.

Éléments à probabilités π_j ($1 \leq j \leq k$) inégales

Définissons d'abord la combinatoire des séries multinomiales. Chaque type j d'élément, parmi les k types possibles, peut apparaître de $n_j = 0$ à n fois dans la série observée, chaque fois avec sa probabilité π_j , ce en observant l'égalité $\sum n_j = n$. Appelons réalisation chaque série $n_1 n_2 \dots n_k$, le nombre total de réalisations étant $Re(n, k) = (n+k-1)!/[n! \times (k-1)!]$.⁴ Chaque réalisation est marquée d'une probabilité distinctive, valable pour chacun de ses $M_k(n; n_j)$ arrangements distincts. Les propriétés globales d'une telle série d'éléments libres, tels le nombre de suites et la longueur de suite maximale, sont associées à et varient selon les $M_k(n; n_j)$ arrangements d'une réalisation.

Le nombre total de suites, S

Pour obtenir la distribution de S , le nombre total de suites dans une série à éléments multiples libres, il suffit de faire la convolution de la fonction récurrente (19) avec la probabilité multinomiale (9), à travers toutes les réalisations $n_1 n_2 \dots n_k$ admissibles. Il s'agit là d'un travail considérable, même sur ordinateur, et la multiplicité des configurations possibles rend utopique la préparation de tables de valeurs critiques. MOOD (1940) et DAVID et BARTON (1962) discutent d'approximations; d'après ces derniers, l'approximation par une binomiale ajustée aux deux premiers moments de S (voir plus bas) donne les meilleurs résultats. Le recours aux méthodes Monte Carlo est un autre expédient, sans doute plus précis.

Le nombre total de suites (S) dans une série de n éléments à k types, avec probabilités $\pi_1 \pi_2 \dots \pi_k$, varie

entre 1 et n . Ses deux premiers moments sont :

$$\mu(S|n, \pi_1 \pi_2 \dots \pi_k) = 1 + (n - 1) \times (1 - \sum_{j=1}^k \pi_j^2) \quad (25a)$$

$$\sigma^2(S|n, \pi_1 \pi_2 \dots \pi_k) = (n - 1) \sum \pi_j^2 + 2(n - 2) \sum \pi_j^3 - (3n - 5) \left[\sum \pi_j^2 \right]^2. \quad (25b)$$

L'espérance (25a) peut être démontrée par un argument récursif. Posons $E(S_n) = \mu(S|n, \pi_1 \pi_2 \dots \pi_k)$ et, par définition, $S_1 = 1$. Utilisant $E(S_n) = E(S_n - 1 + I_n) = E(S_n - 1) + E(I_n)$, nous voyons que le n^e élément ajoutera $I_n = 1$ s'il diffère de l'élément qui le précède, et 0 sinon. Le n^e élément peut être de type 1 avec probabilité π_1 , et il différera du précédent avec probabilité $1 - \pi_1$, la probabilité conjointe de cet événement étant $\pi_1(1 - \pi_1)$. Réunissant tous les éléments possibles, nous trouvons $E(I_n) = \sum_{j=1}^k \pi_j(1 - \pi_j) = 1 - \sum_{j=1}^k \pi_j^2$. Nous obtenons donc $E(S_n) = E(S_{n-1}) + (1 - \sum_{j=1}^k \pi_j^2) = E(S_{n-2}) + 2 \times (1 - \sum_{j=1}^k \pi_j^2) = \dots = E(S_1) + (n - 1)(1 - \sum_{j=1}^k \pi_j^2) = 1 + (n - 1) \times (1 - \sum_{j=1}^k \pi_j^2)$, ce qu'il fallait démontrer.

Le nombre de suites pour l'élément a , S_1

Prenant à part les éléments d'un type, désigné a et à probabilité π , et confondant tous les autres en un type collectif, désigné « b » et à probabilité $(1 - \pi)$, les indications déjà fournies pour les suites d'éléments binaires libres s'appliquent ici, y inclus les moments (17a)-(17b) et les valeurs critiques présentées aux tableaux 4 et 5. À toutes fins pratiques, nous reproduisons la covariance $\sigma(S_i, S_j)$ des nombres de suites d'éléments des types i et j (d'après MOOD, 1940), soit :

$$\sigma(S_i, S_j) = -n \times \pi_i \pi_j \times (1 - 2\pi_i - 2\pi_j + 3\pi_i \pi_j) - \pi_i \pi_j \times (2\pi_i + 2\pi_j - 5\pi_i \pi_j). \quad (26)$$

La longueur maximale de suite, L_{max} et (pour l'élément a) L_{max_1}

Sur la longueur maximale de suite de tout type dans un contexte multinomial, rien n'est connu. Seuls, des résultats élémentaires et triviaux sont accessibles, telle la probabilité d'obtenir une seule suite de longueur n , soit $p(L_{max} = n) = \sum_{j=1}^k \pi_j^n$.

Quant à L_{max_1} , la longueur de suite maximale d'un élément choisi, désigné a , on l'obtient en confondant les

4. Rappelons d'abord (voir note 2) le concept de partition d'un nombre, p. ex. les 7 partitions de $n = 5$, soit = 5, 41, 32, 311, 221, 2111, 11111, et ses 2 3-partitions, 311 et 221. Par contraste avec les k -partitions, une k -réalisation du nombre n consiste en sa décomposition en k nombres n_j tels que $0 \leq n_j < n$ et $\sum n_j = n$, incluant leurs permutations. Noter que le nombre $n_j = 0$ est ici permis, sa valeur étant sujette à la probabilité π_j (Lorsque la valeur $n_j = 0$ est interdite, les réalisations restantes sont dénommées compositions.) Ainsi, les $Re(5, 3) = 21$ 3-réalisations du nombre 5 seront : 500, 050, 005, 410, 401, 140, 104, 041, 014, 320, 302, 230, 203, 032, 023, 311, 131, 113, 221, 212, 122. Ce total augmente sérieusement avec n et k , p. ex. $Re(5, k) = 6, 56$ et 252 pour $k = 2, 4$ et 6, comparativement à $Re(10, k) = 11, 286$ et 3003.



$k - 1$ autres types d'éléments en un seul, le premier ayant probabilité π , les autres, une probabilité globale $1 - \pi$, et en traitant la variable comme si elle provenait d'un processus binomial. La distribution de L_{max_1} est alors la convolution des distributions (15a) et (7).

Éléments à probabilités π_j ($1 \leq j \leq k$) égales à $1/k$

Le traitement de séries constituées d'éléments équiprobables, plus fructueux à coup sûr, doit sa facilité au fait que chacune des k^n séries possibles, chaque composition $n_1 n_2 \dots n_k$, chaque arrangement, sont eux-mêmes équiprobables.

Le nombre total de suites, S

Dans un processus multinomial à k types d'éléments équiprobables, le nombre total de suites (S) obéit à une distribution binomiale; en fait, la variable $S - 1$ a la distribution (7) avec les paramètres $n' = n - 1$ et $\pi = (k - 1)/k$. Ses moments sont :

$$\mu(S|n, \pi_j = 1/k) = 1 + (n - 1)(k - 1)/k; \quad (27a)$$

$$\sigma^2(S|n, \pi_j = 1/k) = (n - 1)(k - 1)/k^2 = (\mu - 1)/k; \quad (27b)$$

$$\gamma_1(S) = (2 - k) / \sqrt{(n - 1)(k - 1)}; \quad (27c)$$

$$\gamma^2(S) = (k^2 - 6k + 6) / [(n - 1)(k - 1)]. \quad (27d)$$

Les tableaux 6 à 11 présentent un jeu de valeurs critiques correspondant à différents percentiles P de l'intégrale binomiale, ce pour $k = 2, 3, 4, 5, 6$, et 10 et $n = 2(1)50$. Le nombre de suites relevant d'un type d'éléments (S_1), disons les suites d'éléments a , a été abordé plus haut.

La longueur maximale de suite, L_{max}

L'ignorance, déjà avouée pour les séries multinomiales à probabilités inégales, persiste dans le cas des séries à probabilités égales. Notons néanmoins un résultat curieux touchant l'espérance de L_{max} et que nous avons trouvé par énumération. Cette espérance⁵ a la forme polynomiale :

$$\mu(L_{max}) = 1 + \sum_{i=2}^n c_i / k^{i-1}; \quad (28)$$

les quelques premiers coefficients sont $c_2 = 1, c_3 = k, c_4 = (k - 1)^2 + k, c_5 = (k - 1)^3 + k^2, c_6 = (k - 1)^4 + k^3 - (k - 1), c_7 = (k - 1)^5 + k^4 - (k - 1)(2k - 1)$. La progression de $\mu(L_{max})$ selon n est de type logarithmique.

Les répétitions (R) et les chaînes de 2 éléments (C_2)

Grâce à la relation (5), qui conjugue les répétitions d'éléments (R) et le nombre de chaînes de 2 éléments (C_2) avec la variable $n - S$, complémentaire du nombre total de

suites, les variables R et C_2 partagent une même distribution. Dans un contexte multinomial à k types d'éléments équiprobables, cette distribution est binomiale, de paramètres $n' = n - 1$ et $\pi = 1/k$. À l'image des moments (27a)-(27d) de S , les variables R et C_2 ont pour moments :

$$\mu(R, C_2|n, (\pi_j) = 1/k) = (n - 1)/k \quad (29a)$$

$$\sigma^2(R, C_2|n, (\pi_j) = 1/k) = (n - 1)(k - 1)/k^2 \quad (29b)$$

$$\gamma_1(R, C_2|n, (\pi_j) = 1/k) = (k - 2) / \sqrt{(n - 1)(k - 1)} \quad (29c)$$

$$\gamma_2(R, C_2|n, (\pi_j) = 1/k) = (k^2 - 6k + 6) / [(n - 1)(k - 1)]. \quad (29d)$$

Les tables usuelles de l'intégrale binomiale et leurs jeux de valeurs critiques (p. ex. LAURENCELLE & DUPUIS, 2000) conviennent très bien à ces variables.

Les chaînes (C_r) et suites (L_r) de longueur r : espérances

Dans une série multinomiale à k types équiprobables, le nombre de chaînes de longueur r (C_r) a une espérance facile à démontrer. Soit la fonction $C_r(n)$, dénotant le nombre de chaînes de r éléments dans une série de n , et $\pi = 1/k$. Posant $C_r(1) = C_r(2) = \dots = C_r(r - 1) = 0$, nous avons $C_r(n) = C_r(n - 1) + I_n$, avec $I_n = 1$ si le n^e élément fait accroître le nombre de chaînes, et $I_n = 0$ sinon. Or,

$$\begin{aligned} E\{C_r(n)\} &= E\{C_r(n - 1) + I_n\} \\ &= E\{C_r(n - 1)\} + E\{I_n\} \\ &= E\{C_r(n - 1)\} + \pi^{r-1}, \end{aligned}$$

puisque l'accroissement I_n vaudra 1 si, étant donné le n^e élément, les $r - 1$ éléments qui le précèdent sont du même type, ce qui se produit avec probabilité π^{r-1} . Nous avons donc :

$$\begin{aligned} E\{C_r(n)\} &= E\{C_r(n - 1) + I_n\} \\ &= E\{C_r(n - 1)\} + \pi^{r-1} \\ &= E\{C_r(n - 2)\} + 2\pi^{r-1} \\ &= \dots \\ &= E\{C_r(r - 1)\} + (n - r + 1) \times \pi^{r-1} \end{aligned}$$

$$EC_r(n) = (n - r + 1) \times \pi^{r-1} = (n - r + 1) / k^{r-1}. \quad (30)$$

Chaînes et suites de longueur r sont associées entre elles par les relations (2) et (3). La seconde relation, grâce à la linéarité de l'opérateur d'espérance, permet d'obtenir :

$$E\{L_r(n)\} = \begin{cases} \frac{(n-r)(k-1)^2 + (k^2-1)}{k^{r+1}} & 1 \leq r \leq n - 2 \\ 2(k - 1) / k^{n-1} & r = n - 1 \\ 1 / k^{n-1} & r = n \end{cases} \quad (31)$$

5. Formule conjecturée.



La somme des expressions (31) pour les suites de toutes longueurs r redonne évidemment l'espérance du nombre total de suites, soit (27a).

Les répétats, R'

Chaque série multinomiale de n éléments, à k types équiprobables, a une réalisation $\{n_1 n_2 \dots n_k\}$ quelconque et manifeste $k^* = \sum_{j=1}^k n_j^0$ types présents, avec $1 \leq k^* \leq k$. La série comporte donc $R' = n - k^*$ répétats, une quantité qui varie d'un minimum de $\max(0, n - k^*)$ à un maximum de $n - 1$. Considérant les k types multinomiaux comme des cellules et les n éléments comme des balles, nous constatons être en présence de la distribution dite d'occupation (DAVID & BARTON, 1962; JOHNSON, KOTZ & KEMP, 1992), k^* représentant le nombre de cellules occupées, $k - k^*$ le nombre de cellules vides et $n - k^* = R'$, le nombre de balles en surplus dans les cellules. Le problème des répétats, banal lorsque le ratio $k : n$ approche zéro, prend de l'intérêt lorsque $k \rightarrow n$.

La distribution de probabilités de R' , à l'instar du nombre de cellules occupées après le lancement de n balles vers k cellules, a pour forme :

$$p(R') = \binom{k}{n - R'} \sum_{j=0}^{n-R'-1} (-1)^j \binom{n - R'}{j} \left(\frac{n - R' - j}{k}\right) \quad (32)$$

Les deux premiers moments de R' sont :

$$\mu(R') = n - k + \frac{(k - 1)^n}{k^{n-1}}; \quad (33a)$$

$$\sigma^2(R') = \frac{(k - 1)(k - 2)^n + (k - 1)^n}{k^{n-1}} - \frac{(k - 1)^{2n}}{k^{2(n-1)}}. \quad (33b)$$

L'espérance (33a) peut être démontrée par l'argument suivant. Soit $\pi_1 = 1/k$, la probabilité qu'apparaisse un élément d'un type donné, disons a , et R'_1 , le nombre de répétats de a . Dans une série de n apparitions, la probabilité d'obtenir t éléments a est binomiale, soit $p(t) = \binom{n}{t} \pi_1^t (1 - \pi_1)^{n-t}$. Comme, pour cet élément, $R'_1 = t - 1$, son espérance est : $ER'_1 = \sum_{t \geq 1} (t - 1) \cdot p(t) = \sum_{t \geq 1} t \cdot p(t) - \sum_{t \geq 1} p(t) = n \times \pi_1 - [1 - (1 - \pi_1)^n]$.

De là, pour le nombre total de répétats R' , où $R' = R'_1 + R'_2 + \dots + R'_k$ et des séries à types d'éléments équiprobables, i.e. $\pi_j = \pi_1 = 1/k$, tout j , nous obtenons : $E\{R'\} = k\{n \times \pi_1 - [1 - (1 - \pi_1)^n]\} = n - k + (k - 1)^n / k^{n-1}$, comme prévu.

L'espérance du nombre de répétats est une fonction de k et n , disons $R'(k, n)$. Elle tend vers $n - k$ pour n croissant et vers 0 pour k croissant. Quant à l'espérance $R'(n, n)$, elle avoisine $n \times e^{-1} \approx 0,37n$ (où $e \approx 2,7183$) pour n raisonnablement fort.

Que ce soit pour des études d'échantillonnage, des jeux de hasard ou dans d'autres contextes, il est intéressant de considérer à part les séries ne contenant aucun répétat, i.e. celles pour lesquelles $R' = 0$. Par (32), nous avons $p(R' = 0) = 0$ pour $n > k$. Le tableau 12 présente quelques valeurs de $p(R' = 0)$.

Un exemple

Prenons comme exemple un processus multinomial à probabilités inégales, soit le processus $M_3(n = 12 | \pi_j = 0, 5; 0, 3; 0, 2)$. On nous propose la série suivante :

$$\{b \ a \ c \ c \ c \ a \ a \ a \ a \ a \ b \ c\},$$

La série, avec $n = 12$ éléments, présente $S = 6$ suites (dans un intervalle de 1 à 12), une longueur de suite maximale $L_{max} = 5$ (dans un intervalle de 1 à 12), et $R' = 9$ répétats (dans un intervalle de 9 à 11). La série observée manifeste-t-elle des propriétés d'ordre remarquables ?

Pour connaître les distributions des différentes statistiques (S, L_{max}, R') à l'étude, nous devons recourir ici, soit à un algorithme d'énumération, soit à un échantillonnage des séries possibles par la méthode Monte Carlo. Nous avons procédé par énumération.

Il y a $Re(12, 3) = 91$ réalisations possibles lesquelles, en permutant leurs éléments, peuvent générer 531 441 séries⁶, telle la série générique $\sigma_u \equiv (x_1 \ x_2 \dots \ x_n)_u$, pour $u = 1$ à 531.441, chacune formée au hasard avec la probabilité égale à et caractérisée par une valeur propre des variables S, L_{max} et R' . Le tableau 13 fournit les distributions appropriées de ces variables, obtenues par énumération.

La confrontation des descripteurs de notre série à leurs distributions de probabilités respectives ne laisse voir rien de remarquable. Non seulement le nombre de suites, $S = 6$, est lui-même assez fréquent ($p = 0,12365$) mais le percentile extrême dont il s'approche davantage est d'environ 0,21. De même pour $L_{max} (=5)$, qui correspond ici à une suite d'éléments a , les plus probables, et qui n'atteint que le percentile 0,155. Notre série manifeste le nombre minimal de répétats, i.e. $n - k$, de forte probabilité.

Pour éclairer le lecteur, nous avons mis en œuvre trois procédés d'approximation de la distribution $p(S)$ par une binomiale, tel que suggéré par DAVID et BARTON (1962). Le procédé le plus simple (B_1) consiste à « ramener » les k types à des valeurs de probabilité égales, la variable $S - 1$ revenant alors à une binomiale à paramètres $n' = n - 1$ et $\pi = (k - 1)/k$. Un autre procédé consiste cette fois à estimer des paramètres convenables de la binomiale, en utilisant les premiers moments de S . Par le tableau 6 ou nos formules (25a) et (25b), nous avons ici $\mu(S) = 7,82$ et $\sigma^2(S) = 2,90360$. Nous estimons alors :

6. Nombre obtenu sur ordinateur, par énumération et permutation des réalisations.



$\hat{\pi} = 1 - \sigma^2(S)/[\mu(S) - 1]$ et $\hat{n}' = (\mu - 1)/\hat{\pi}$, obtenant $\hat{\pi} \approx 0,57425$, puis $\hat{n}' \approx 11,87632$. Pour simplifier le calcul de la probabilité binomiale, nous pouvons ramener \hat{n}' à une valeur entière, disons $\hat{n}' = 12$, puis reprendre $\hat{\pi} = (\mu(S) - 1)/\hat{n}' \approx 0,56833$, obtenant un procédé (B_2) de calcul facile. Ou bien, nous conservons les valeurs estimatives de \hat{n}' de $\hat{\pi}$ et calculons la probabilité binomiale avec des coefficients non entiers, à l'aide de la fonction Gamma (voir LAURENCELLE & DUPUIS, 2000) : c'est le procédé B_3 . La comparaison des distributions de probabilités, faite à l'aide de la différence maximale de Kolmogorov-Smirnov, i.e. $d_{max} = |Pr_1(S \leq r) - Pr_2(S \leq r)|_{\text{tout } r}$, est nettement à l'avantage du procédé plus précis, et moins commode. En effet, en comparant les trois distributions approximatives à la distribution exacte de S , donnée au tableau 6, nous obtenons $d_{max}(B_1) \approx 0,1222$, $d_{max}(B_2) \approx 0,1428$ et $d_{max}(B_3) \approx 0,0092$.

Références

BRADLEY, J. V. (1968). *Distribution-free statistical tests*. Prentice-Hall : Englewood Cliffs (NJ).
 DAVID, F. N. & BARTON, D. E. (1962). *Combinatorial chance*. Charles Griffin : London.
 JOHNSON, N. L., KOTZ, S. & KEMP, A. W. (1992). *Univariate discrete distributions (2e édition)*. Wiley : New York.

LAURENCELLE, L. (1986). *Observer le réel : quelques questions d'intérêt méthodologique*. M. Lirette, et M. Piéron (dir.), *Méthodologie de la recherche en enseignement de l'activité physique et des sports*. DSAP, UQTR : in C. Paré.
 LAURENCELLE, L. (1995). Le nombre total de suites : moments, approximations et valeurs critiques. In *Actes du colloque sur les Méthodes et Applications de la Statistique* (p. 73-88). Sainte-Foy : Bureau de la statistique du Québec.
 LAURENCELLE, L. & DUPUIS, F. A. (2000). *Tables statistiques expliquées et appliquées (2e édition)*. Le Griffon d'argile : Sainte-Foy.
 MOOD, A. M. (1940). The distribution theory of runs. *Annals of mathematical statistics*, 11, 367-392.
 OWEN, D. B. (1962). *Handbook of statistical tables*. Addison-Wesley : Reading (Mass.)
 SHAUGHNESSY, P. W. (1981). Multiple runs distributions : recurrences and critical values. *Journal of the American Statistical Association*, 76, 732-736.
 SWED, F. S. & EISENHART, C. (1943). Tables for testing the randomness of grouping in a sequence of alternatives. *Annals of mathematical statistics*, 14, 66-87.
 WEISS, L. (1988). Runs. In S. K. N. L. J. (DIR (Éd.), *Encyclopedia of statistical sciences*, vol (p. 222-226). Wiley : 8 (p.. New York.

Encadré 1. Programme d'énumération des arrangements distincts de n éléments répartis en n_1, n_2, \dots, n_k selon leur type (langage de programmation Pascal-Delphi)

```

Program k\_comb {algorithme de combinaisons a k types de donnees, L. Laurencelle, UQTR};
const nmax=20; nmax=30; nhmax=20;
var k,n,j : integer;
    ne : array[1..nmax] of integer;
    X : array[1..nmax] of integer;
    Pos : array[1..nmax] of boolean;
    hist : array[1..nmax] of longint;
procedure analyser {traiter ici l'arrangement obtenu }
    {le traitement illustre compte le nombre total de suites S };
var S,c,i:integer;
begin S:=1; c:=X[1]; i:=1;
    while i<n do
        begin i:=i+1; if X[i]<>X[i - 1] then S:=S+1 end;
        hist[S]:=hist[S]+1
    end;

procedure gen(r,nr:integer);
var i,j,p,dr : integer;
    li : array[1..nmax] of integer;
    v : array[1..nmax] of integer;
label Q,W ;
begin

```



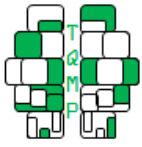

```
dr:=0;
for p:=1 to n do if Pos[p] then begin dr:=dr+1; li[dr]:=p end;
i:=0; v[0]:=0;
repeat i:=i+1; v[i]:=v[i - 1]+1;
  Q : p:=li[v[i]]; Pos[p]:=false; X[p]:=r
until i=nr;
if r<k-1 then gen(r+1,ne[r+1])
  else
  begin
    for j:=1 to n do if Pos[j] then X[j]:=k;
    analyser
  end;
W: Pos[p]:=true;
  if v[i]<dr-nr+i then begin v[i]:=v[i]+1; goto Q end
  else begin i:=i-1; if i>0 then
    begin p:=li[v[i]]; goto W end
  end
end;

begin {corps du programme}
  write ('k_:_'); readln(k); n:=0;

  for j:=1 to k do begin write (j:3, ' '); readln(ne[j]); n:=n+ne[j]
  end;
  for j:=1 to n do Pos[j]:=true;
  for j:=k to n do hist[j]:=0;
  gen(1,ne[1]); writeln(' _S_ f(S) ');
  for j:=k to n do if hist[j]>0 then writeln(j:3,hist[j]:10)
end.
```

Encadré 2. Programme de calcul par récursion du nombre d'arrangements distincts de n éléments répartis en n_1, n_2, \dots, n_k selon leur type et présentant S suites d'éléments semblables (langage de programmation Pascal-Delphi)

```
program saughnessy
  {recurrence pour le nombre de suites d'elements multiples, d'apres P.W. Saughnessy,
  JASA, 1981, vol. 76, p. 732-736, L. Laurencelle };
var k,n,i,S:integer;
    ne,nt:array[1..20] of integer;
    tot,f:double;
    ok:boolean;
function fact(x:integer):double; {fonction Factorielle}
var i:integer; f:double;
begin
  if x<0 then fact:=0 else if x<2 then fact:=1 else
  begin f:=x; for i:=2 to x-1 do f:=f*i; fact:=f end
end;
function comb(a,b:integer):double; {fonction Combinaisons}
var x:integer; nu,de:double;
begin
  if a<b then comb:=0
  else if b=0 then comb:=1
```



```
else if b<0 then comb:=0 else
begin nu:=a; de:=b;
  for x:=1 to b-1 do begin nu:=nu*(a-x); de:=de*(b-x) end;
  comb:=nu / de
end
end;

function recR(VR,v:integer):double; {Recurrence 1 de Saugnessy}
var r,s,j,nnv:integer; t,xR,f:double;
begin r:=VR-v;nnv:=nt[v]-ne[v]-v; t:=0;
  for s:=1 to r+1 do
    begin if v>2 then xR:=recR(VR-s,v-1) else
      if VR-s=1 then xR:=1 else xR:=0;
    if xR>0 then
      begin f:=0; for j:=0 to nnv+1 do
        f:=f+comb(nnv-r+s,j)*comb(ne[v]-1,s-1-j)*comb(VR-s+1,s-2*j)
      end;
      t:=t+xR*f;
    end;
  recR:=t
end;
begin {Corps du programme}
write(' k: ');readln(k); nt[0]:=0; tot:=1;
for i:=1 to k do
  begin write(i:2,' ');readln(ne[i]);
  nt[i]:=nt[i-1]+ne[i]; tot:=tot*fact(ne[i]) end;
n:=nt[k]; tot:=fact(n)/tot; S:=k;ok:=true;
while ok do begin f:=recR(S,k);
  if f>0 then writeln(S:4,f/tot:10:5,f:18:0)
  else ok:=false;
  if ok then if S<n then S:=S+1 else ok:=false
end
end.
```

Citation

LAURENCELLE, L. (2020). Une étude statistique des séries de deux ou plusieurs types d'éléments The statistical study of series of two or many types of elements : a digest. *The Quantitative Methods for Psychology*, 16(4), 391-417. doi :10.20982/tqmp.16.4.p391

Copyright © 2020, LAURENCELLE. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 12/05/2020 ~ Accepted: 12/06/2020

Les tableaux suivent.



Tableau 1 ■ Table de significativité de L_{max} au seuil de 0,05, soit la longueur maximale d’une suite d’éléments semblables dans une série comportant n_1 et n_2 éléments de deux types : le nombre indiqué est la valeur maximale de n_1 pour laquelle L_{max} atteint le seuil de probabilité, selon $d = n_2 - n_1$ et n_1 . †

Seuil de 0,05								
d \ Lmax	5	6	7	8	9	10	11	12
+ 0	(5)5	7	11	17	27	45	76	> 100
+ 1	-	(5)6	10	16	27	44	75	> 100
+ 2	-	(5)5	9	15	25	43	74	> 100
+ 3	-	-	(5)7	13	24	42	73	> 100
+ 4	-	-	-	(5)10	22	40	72	> 100
+ 5	-	-	-	(5)6	19	38	70	> 100
+ 6	-	-	-	-	(5)14	35	68	> 100
+ 7	-	-	-	-	-	(5)31	66	> 100
+ 8	-	-	-	-	-	(5)26	63	> 100
+ 9	-	-	-	-	-	-	(5)60	> 100
+ 10	-	-	-	-	-	-	(5)56	> 100

Seuil de 0,01								
d \ Lmax	7	8	9	10	11	12	13	14
+ 0	(7)7	10	14	20	30	46	73	>100
+ 1	(6)6	9	13	19	29	45	72	>100
+ 2	(5)5	8	12	18	28	44	71	>100
+ 3	-	(5)6	10	16	26	42	69	>100
+ 4	-	-	(5)7	13	23	40	68	>100
+ 5	-	-	(4)4	9	20	37	66	>100
+ 6	-	-	-	(4)5	17	34	63	>100
+ 7	-	-	-	-	(4)9	29	60	>100
+ 8	-	-	-	-	-	(4)23	56	>100
+ 9	-	-	-	-	-	(3)3	51	>100
+ 10	-	-	-	-	-	-	(3)45	>100

Seuil de 0,001									
d \ Lmax	9	10	11	12	13	14	15	16	17
+ 0	(9)9	12	15	20	26	36	52	77	>100
+ 1	(8)8	11	14	18	26	36	51	76	>100
+ 2	(7)7	9	13	17	24	34	50	75	>100
+ 3	-	(7)7	11	15	22	32	48	73	>100
+ 4	-	(6)6	8	12	19	29	45	71	>100
+ 5	-	-	(6)6	10	16	26	42	68	>100
+ 6	-	-	-	(6)7	12	22	38	65	>100
+ 7	-	-	-	(5)5	8	16	33	61	>100
+ 8	-	-	-	-	(5)5	11	27	56	>100
+ 9	-	-	-	-	-	(5)6	19	50	99
+ 10	-	-	-	-	-	-	(5)8	42	94

Note. † : Le nombre entre parenthèses, au début de chaque ligne, indique la valeur minimale de n_1 requise pour que L_{max} soit significatif à un seuil donné, sous n_1 et $n_2 = n_1 + d$.

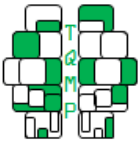


Tableau 2 Valeurs critiques du nombre total de suites (S) d'une série de Bernoulli à paramètre π, aux seuils de signification ,005, ,01, ,025 et ,05

Table with 50 rows (n=4 to 50) and 20 columns (π = 1/2 to π = 1/10). Each cell contains a numerical value representing the critical number of sequences.



Tableau 3 Valeurs critiques du nombre total de suites (S) d'une série de n essais de Bernoulli à paramètre π, aux seuils de signification 0,95, 0,975, 0,99 et 0,995

Table with 50 rows (n) and columns grouped by alpha (π) values: 1/2, 1/3, 1/5, 1/4, 1/6, 1/10. Each column contains critical values for alpha = 0.95, 0.975, 0.99, 0.995.

Tableau 4 ■ Valeurs critiques du nombre de suites (S_1) de l'élément α dans une série de n essais de Bernoulli de paramètre $\pi (= p(\alpha))$, aux seuils de signification 0,005, 0,01, 0,025 et 0,05

n	$\pi = 1/2$				$\pi = 1/3$				$\pi = 1/5$				$\pi = 1/4$				$\pi = 1/6$				$\pi = 1/10$										
	,005	,01	,025	,05	,005	,01	,025	,05	,005	,01	,025	,05	,005	,01	,025	,05	,005	,01	,025	,05	,005	,01	,025	,05	,005	,01	,025	,05	,005	,01	,025
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
5	-	-	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
6	-	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
7	-	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
8	0	0	0	0	-	-	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
9	0	0	0	0	-	-	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
10	0	0	0	0	-	-	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
11	0	0	0	1	-	-	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
12	0	0	1	1	-	-	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
13	0	0	1	1	-	-	0	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
14	0	1	1	1	0	0	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
15	1	1	1	1	0	0	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
16	1	1	1	2	0	0	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
17	1	1	2	2	0	0	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
18	1	1	2	2	0	0	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
19	1	2	2	2	0	0	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
20	1	2	2	3	0	0	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
21	2	2	2	3	0	0	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
22	2	2	2	3	1	1	1	2	0	0	0	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
23	2	2	3	3	1	1	1	2	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
24	2	2	3	3	1	1	2	2	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
25	2	3	3	3	1	1	2	2	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
26	3	3	3	4	1	1	2	2	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
27	3	3	3	4	1	1	2	2	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
28	3	3	4	4	2	2	2	2	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
29	3	3	4	4	2	2	2	2	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
30	3	4	4	5	2	2	2	2	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
31	3	4	4	5	2	2	2	2	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
32	4	4	4	5	2	2	2	3	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
33	4	4	5	5	2	2	2	3	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
34	4	4	5	5	2	2	2	3	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
35	4	5	5	6	3	3	3	4	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
36	4	5	5	6	3	3	3	4	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
37	5	5	6	6	3	3	3	4	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
38	5	5	6	6	3	3	3	4	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
39	5	5	6	6	3	3	4	4	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
40	5	6	6	7	4	4	4	5	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
41	5	6	6	7	4	4	4	5	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
42	6	6	6	7	4	4	4	5	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
43	6	6	7	7	4	4	4	5	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
44	6	6	7	8	4	4	4	5	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
45	6	7	7	8	4	4	5	6	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
46	6	7	7	8	5	5	6	6	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
47	7	7	7	8	5	5	6	6	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
48	7	7	8	8	5	5	6	7	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
49	7	7	8	9	5	5	6	7	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
50	7	8	8	9	5	5	6	7	0	0	0	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	



Tableau 5 Valeurs critiques du nombre de suites (S1) de l'élément α dans une série de n essais de Bernoulli de paramètre π (= p(α)), aux seuils de signification 0,95, 0,975, 0,99 et 0,995

Table with columns for n and π = 1/2, 1/3, 1/5, 1/4, 1/6, 1/10. Each column contains values for significance levels .95, .975, .99, .995.



Tableau 6 ■ Valeurs critiques du nombre total de suites (S) pour des séries de n éléments tirées d'un processus 2-(ou bi)nomial à k types équiprobables.

n / P	,005	,01	,025	,05	,95	,975	,99	,995
2	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-
6	-	-	-	1	6	-	-	-
7	-	-	1	1	7	7	-	-
8	-	1	1	1	8	8	8	-
9	1	1	1	2	8	9	9	9
10	1	1	2	2	9	9	10	10
11	1	1	2	2	10	10	11	11
12	1	2	2	3	10	11	11	12
13	2	2	3	3	11	11	12	12
14	2	2	3	4	11	12	13	13
15	2	3	3	4	12	13	13	14
16	3	3	4	4	13	13	14	14
17	3	3	4	5	13	14	15	15
18	3	4	5	5	14	14	15	16
19	4	4	5	6	14	15	16	16
20	4	5	5	6	15	16	16	17
21	4	5	6	6	16	16	17	18
22	5	5	6	7	16	17	18	18
23	5	6	6	7	17	18	18	19
24	5	6	7	8	17	18	19	20
25	6	6	7	8	18	19	20	20
26	6	7	8	8	19	19	20	21
27	7	7	8	9	19	20	21	21
28	7	8	8	9	20	21	21	22
29	7	8	9	10	20	21	22	23
30	8	8	9	10	21	22	23	23
31	8	9	10	11	21	22	23	24
32	8	9	10	11	22	23	24	25
33	9	9	10	11	23	24	25	25
34	9	10	11	12	23	24	25	26
35	10	10	11	12	24	25	26	26
36	10	11	12	13	24	25	26	27
37	10	11	12	13	25	26	27	28
38	11	11	13	14	25	26	28	28
39	11	12	13	14	26	27	28	29
40	12	12	13	14	27	28	29	29
41	12	13	14	15	27	28	29	30
42	12	13	14	15	28	29	30	31
43	13	14	15	16	28	29	30	31
44	13	14	15	16	29	30	31	32
45	14	14	16	17	29	30	32	32
46	14	15	16	17	30	31	32	33
47	14	15	16	17	31	32	33	34
48	15	16	17	18	31	32	33	34
49	15	16	17	18	32	33	34	35
50	16	16	18	19	32	33	35	35

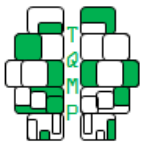


Tableau 7 ■ Valeurs critiques du nombre total de suites (S) pour des séries de n éléments tirées d'un processus 3-nomial à k types équiprobables.

n / P	,005	,01	,025	,05	,95	,975	,99	,995
2	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-
4	-	-	-	1	-	-	-	-
5	-	-	1	1	-	-	-	-
6	1	1	1	2	-	-	-	-
7	1	1	2	2	-	-	-	-
8	1	2	2	3	-	-	-	-
9	2	2	3	3	9	-	-	-
10	2	3	3	4	10	-	-	-
11	3	3	4	4	11	11	-	-
12	3	4	4	5	12	12	-	-
13	4	4	5	5	13	13	13	-
14	4	5	5	6	13	14	14	-
15	5	5	6	6	14	15	15	15
16	5	6	6	7	15	15	16	16
17	6	6	7	8	16	16	17	17
18	6	7	7	8	16	17	17	18
19	7	7	8	9	17	18	18	19
20	7	8	9	9	18	18	19	19
21	8	8	9	10	19	19	20	20
22	8	9	10	10	19	20	21	21
23	9	9	10	11	20	21	21	22
24	9	10	11	12	21	22	22	23
25	10	10	11	12	22	22	23	23
26	10	11	12	13	22	23	24	24
27	11	12	13	13	23	24	25	25
28	11	12	13	14	24	25	25	26
29	12	13	14	14	25	25	26	27
30	13	13	14	15	25	26	27	27
31	13	14	15	16	26	27	28	28
32	14	14	15	16	27	28	28	29
33	14	15	16	17	28	28	29	30
34	15	16	17	17	28	29	30	31
35	15	16	17	18	29	30	31	31
36	16	17	18	19	30	31	31	32
37	17	17	18	19	31	31	32	33
38	17	18	19	20	31	32	33	34
39	18	18	20	20	32	33	34	34
40	18	19	20	21	33	34	35	35
41	19	20	21	22	33	34	35	36
42	19	20	21	22	34	35	36	37
43	20	21	22	23	35	36	37	37
44	20	21	22	24	36	37	38	38
45	21	22	23	24	36	37	38	39
46	22	22	24	25	37	38	39	40
47	22	23	24	25	38	39	40	40
48	23	24	25	26	39	39	41	41
49	23	24	25	27	39	40	41	42
50	24	25	26	27	40	41	42	43



Tableau 8 ■ Valeurs critiques du nombre total de suites (S) pour des séries de n éléments tirées d'un processus 4-nomial à k types équiprobables.

n / P	,005	,01	,025	,05	,95	,975	,99	,995
2	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-
4	-	-	1	1	-	-	-	-
5	1	1	1	1	-	-	-	-
6	1	1	2	2	-	-	-	-
7	2	2	2	3	-	-	-	-
8	2	2	3	3	-	-	-	-
9	3	3	3	4	-	-	-	-
10	3	4	4	5	-	-	-	-
11	4	4	5	5	-	-	-	-
12	4	5	5	6	12	-	-	-
13	5	5	6	6	13	-	-	-
14	5	6	7	7	14	14	-	-
15	6	6	7	8	15	15	-	-
16	7	7	8	8	16	16	-	-
17	7	8	8	9	17	17	-	-
18	8	8	9	10	18	18	18	-
19	8	9	10	10	18	19	19	-
20	9	10	10	11	19	20	20	20
21	10	10	11	12	20	20	21	21
22	10	11	12	12	21	21	22	22
23	11	12	12	13	22	22	23	23
24	12	12	13	14	22	23	24	24
25	12	13	14	14	23	24	24	25
26	13	13	14	15	24	25	25	26
27	13	14	15	16	25	26	26	27
28	14	15	16	16	26	26	27	27
29	15	15	16	17	27	27	28	28
30	15	16	17	18	27	28	29	29
31	16	17	18	18	28	29	30	30
32	17	17	18	19	29	30	30	31
33	17	18	19	20	30	31	31	32
34	18	19	20	21	31	31	32	33
35	19	19	20	21	31	32	33	33
36	19	20	21	22	32	33	34	34
37	20	21	22	23	33	34	35	35
38	21	21	22	23	34	35	35	36
39	21	22	23	24	35	35	36	37
40	22	23	24	25	36	36	37	38
41	23	23	24	25	36	37	38	38
42	23	24	25	26	37	38	39	39
43	24	25	26	27	38	39	40	40
44	25	25	26	27	39	40	40	41
45	25	26	27	28	40	40	41	42
46	26	27	28	29	40	41	42	43
47	27	27	29	30	41	42	43	43
48	27	28	29	30	42	43	44	44
49	28	29	30	31	43	44	45	45
50	29	29	31	32	44	44	45	46



Tableau 9 ■ Valeurs critiques du nombre total de suites (S) pour des séries de n éléments tirées d'un processus 5-nomial à k types équiprobables.

n / P	,005	,01	,025	,05	,95	,975	,99	,995
2	-	-	-	-	-	-	-	-
3	-	-	-	1	-	-	-	-
4	-	1	1	1	-	-	-	-
5	1	1	1	2	-	-	-	-
6	1	2	2	2	-	-	-	-
7	2	2	3	3	-	-	-	-
8	3	3	3	4	-	-	-	-
9	3	3	4	4	-	-	-	-
10	4	4	5	5	-	-	-	-
11	4	5	5	6	-	-	-	-
12	5	5	6	6	-	-	-	-
13	6	6	7	7	-	-	-	-
14	6	7	7	8	-	-	-	-
15	7	7	8	9	15	-	-	-
16	8	8	9	9	16	-	-	-
17	8	9	9	10	17	-	-	-
18	9	9	10	11	18	18	-	-
19	10	10	11	11	19	19	-	-
20	10	11	12	12	20	20	-	-
21	11	12	12	13	21	21	-	-
22	12	12	13	14	22	22	22	-
23	12	13	14	14	22	23	23	-
24	13	14	14	15	23	24	24	-
25	14	14	15	16	24	25	25	25
26	14	15	16	17	25	26	26	26
27	15	16	17	17	26	26	27	27
28	16	16	17	18	27	27	28	28
29	17	17	18	19	28	28	29	29
30	17	18	19	20	29	29	30	30
31	18	19	19	20	29	30	31	31
32	19	19	20	21	30	31	31	32
33	19	20	21	22	31	32	32	33
34	20	21	22	22	32	33	33	34
35	21	21	22	23	33	33	34	34
36	21	22	23	24	34	34	35	35
37	22	23	24	25	35	35	36	36
38	23	24	25	25	35	36	37	37
39	24	24	25	26	36	37	38	38
40	24	25	26	27	37	38	38	39
41	25	26	27	28	38	39	39	40
42	26	26	28	28	39	39	40	41
43	26	27	28	29	40	40	41	42
44	27	28	29	30	41	41	42	42
45	28	29	30	31	41	42	43	43
46	29	29	31	31	42	43	44	44
47	29	30	31	32	43	44	45	45
48	30	31	32	33	44	45	45	46
49	31	32	33	34	45	45	46	47
50	32	32	33	34	46	46	47	48



Tableau 10 ■ Valeurs critiques du nombre total de suites (S) pour des séries de n éléments tirées d'un processus 6-nomial à k types équiprobables.

n / P	,005	,01	,025	,05	,95	,975	,99	,995
2	-	-	-	-	-	-	-	-
3	-	-	-	1	-	-	-	-
4	1	1	1	1	-	-	-	-
5	1	1	2	2	-	-	-	-
6	2	2	2	3	-	-	-	-
7	2	3	3	3	-	-	-	-
8	3	3	4	4	-	-	-	-
9	4	4	4	5	-	-	-	-
10	4	5	5	6	-	-	-	-
11	5	5	6	6	-	-	-	-
12	6	6	7	7	-	-	-	-
13	6	7	7	8	-	-	-	-
14	7	7	8	8	-	-	-	-
15	8	8	9	9	-	-	-	-
16	8	9	9	10	-	-	-	-
17	9	9	10	11	-	-	-	-
18	10	10	11	11	18	-	-	-
19	10	11	12	12	19	-	-	-
20	11	12	12	13	20	-	-	-
21	12	12	13	14	21	-	-	-
22	13	13	14	15	22	22	-	-
23	13	14	15	15	23	23	-	-
24	14	15	15	16	24	24	-	-
25	15	15	16	17	25	25	-	-
26	16	16	17	18	26	26	-	-
27	16	17	18	18	27	27	27	-
28	17	18	18	19	27	28	28	-
29	18	18	19	20	28	29	29	-
30	18	19	20	21	29	30	30	-
31	19	20	21	21	30	31	31	31
32	20	21	22	22	31	32	32	32
33	21	21	22	23	32	32	33	33
34	21	22	23	24	33	33	34	34
35	22	23	24	25	34	34	35	35
36	23	24	25	25	35	35	36	36
37	24	24	25	26	35	36	37	37
38	24	25	26	27	36	37	37	38
39	25	26	27	28	37	38	38	39
40	26	27	28	29	38	39	39	40
41	27	27	28	29	39	40	40	41
42	28	28	29	30	40	40	41	42
43	28	29	30	31	41	41	42	42
44	29	30	31	32	42	42	43	43
45	30	31	32	32	43	43	44	44
46	31	31	32	33	43	44	45	45
31	32	33	34	44	45	46	46	47
48	32	33	34	35	45	46	47	47
49	33	34	35	36	46	47	47	48
50	34	34	35	36	47	48	48	49



Tableau 11 ■ Valeurs critiques du nombre total de suites (S) pour des séries de n éléments tirées d'un processus 10-nomial à k types équiprobables.

n / P	,005	,01	,025	,05	,95	,975	,99	,995
2	-	-	-	-	-	-	-	-
3	-	1	1	1	-	-	-	-
4	1	1	1	2	-	-	-	-
5	2	2	2	2	-	-	-	-
6	2	3	3	3	-	-	-	-
7	3	3	4	4	-	-	-	-
8	4	4	4	5	-	-	-	-
9	4	5	5	6	-	-	-	-
10	5	6	6	6	-	-	-	-
11	6	6	7	7	-	-	-	-
12	7	7	8	8	-	-	-	-
13	8	8	8	9	-	-	-	-
14	8	9	9	10	-	-	-	-
15	9	10	10	11	-	-	-	-
16	10	10	11	11	-	-	-	-
17	11	11	12	12	-	-	-	-
18	12	12	13	13	-	-	-	-
19	12	13	13	14	-	-	-	-
20	13	14	14	15	-	-	-	-
21	14	14	15	16	-	-	-	-
22	15	15	16	16	-	-	-	-
23	16	16	17	17	-	-	-	-
24	16	17	18	18	-	-	-	-
25	17	18	18	19	-	-	-	-
26	18	19	19	20	-	-	-	-
27	19	19	20	21	-	-	-	-
28	20	20	21	22	-	-	-	-
29	21	21	22	22	-	-	-	-
30	21	22	23	23	30	-	-	-
31	22	23	23	24	31	-	-	-
32	23	24	24	25	32	-	-	-
33	24	24	25	26	33	-	-	-
34	25	25	26	27	34	-	-	-
35	25	26	27	28	35	-	-	-
36	26	27	28	28	36	-	-	-
37	27	28	29	29	37	37	-	-
38	28	29	29	30	38	38	-	-
39	29	29	30	31	39	39	-	-
40	30	30	31	32	40	40	-	-
41	30	31	32	33	41	41	-	-
42	31	32	33	34	42	42	-	-
43	32	33	34	34	43	43	-	-
44	33	34	35	35	44	44	-	-
45	34	34	35	36	45	45	45	-
46	35	35	36	37	46	46	46	-
47	36	36	37	38	46	47	47	-
48	36	37	38	39	47	48	48	-
49	37	38	39	40	48	49	49	-
50	38	39	40	40	49	50	50	-



Tableau 12 ■ Probabilité de n’observer aucun répétat ($R' = 0$) dans des séries multinomiales de n éléments à k types équiprobables.

$n \backslash k$	2	3	4	5	6	7	8	9	10
2	,5	,6667	,75	,8	,8333	,8571	,875	,8889	,9
3	-	,2222	,375	,48	,5556	,6122	,6563	,6914	,72
4	-	-	,0938	,192	,2778	,3499	,4102	,4609	,504
5	-	-	-	,0384	,0926	,1499	,2051	,2561	,3024
6	-	-	-	-	,0154	,0428	,0769	,1138	,1512
7	-	-	-	-	-	,0061	,0192	,0379	,0605
8	-	-	-	-	-	-	,0024	,0084	,0181
9	-	-	-	-	-	-	-	,0009	,0036
10	-	-	-	-	-	-	-	-	,0004

Tableau 13 ■ Distributions de probabilités des statistiques S , L_{max} et R' dans des séries de type $M_3(12|0,5; 0,3; 0,2)$ à π_j inégales

S	$p(S)$	L_{max}	$p(L_{max})$	R'	$p(R')$
1	,0002	12	,0002	11	,0002
2	,0011	11	,0005	10	,0823
3	,0067	10	,0012	9	,9174
4	,0214	9	,003		
5	,0614	8	,007		
6	,1237	7	,0164		
7	,1967	6	,0383		
8	,2301	5	,0886		
9	,1965	4	,1986		
10	,1145	3	,3679		
11	,041	2	,2715		
12	,0068	1	,0068		