

La loi de Pascal restreinte et ses cas particuliers

Constrained Pascal distribution and its special cases

Louis Laurencelle

Université du Québec à Trois-Rivières

Pascal distribution, $Pa(r, \pi)$, also called Negative binomial distribution, pertains to the trial number n at which the first r successes have been obtained, each trial being the realization of a Bernoulli variable with success probability π . We review basic concepts of the Bernoulli process and delve into the “constrained Pascal” distribution, or Negative binomial distribution of order k , $Pa(r, k, \pi)$, which concerns the trial number n at which r successes have been recorded *within the last k trials* for the first time. Practical functions and procedures are given for producing probability mass and distribution functions, along with indications for estimating the π parameter.

*Partiellement basé sur Lettres Statistiques, 1983, vol. 7, p. 1.1-1.25; 1987, vol. 8, p. 25-47 ;
 1998, vol. 10, p. 51-57 ; 2005, vol. 12, p. 65-79, avec des compléments*

Imaginons une suite d'événements qui se produisent ou qu'on peut invoquer à volonté et dont l'apparition présente l'un de deux résultats aléatoires, par exemple réussite (R) ou échec (E). Si les événements successifs, tout comme leur résultat, se produisent indépendamment les uns des autres, et que la probabilité de chaque résultat est constante, nous avons affaire à un processus de Bernoulli, lequel a servi et sert de modèle dans bien des situations pratiques, en physique des particules, en statistique ainsi que, bien entendu, dans le calcul des jeux de hasard. Le jeu de Pile ou Face en est l'exemple par excellence.

La variable de Bernoulli, disons X , peut prendre l'une de deux valeurs, soit :

$X = 1$: Succès, tourner à droite, maladie contractée, face 1 d'un dé, etc., ou

$X = 0$: Échec, tourner à gauche, maladie évitée, faces 2 à 6 d'un dé, etc.

La réalisation d'une suite d'événements pourra prendre la forme de l'exemple ci-dessous :

(Exemple A)

Événement	1	2	3	4	5	6	7	8	9	...
Résultat	1	0	1	0	0	1	1	0	1	...

Les probabilités, constantes d'un événement (ou essai) à l'autre, sont $\pi = \Pr\{X = 1\}$ et $\omega = 1 - \pi = \Pr\{X = 0\}$. Ainsi, la

probabilité d'observer deux succès consécutifs en deux essais sera $\Pr\{X_1 = 1\} \times \Pr\{X_2 = 1\} = \pi \cdot \pi = \pi^2$.

La loi binomiale. Le processus de Bernoulli se prête à maintes structurations mathématiques. La plus connue est celle de la loi binomiale, laquelle consiste à étudier la distribution du nombre de succès obtenus dans une suite finie d'essais. Soit $n, n = 1, 2, \dots$, le nombre d'essais de Bernoulli effectués, et $r = \sum X_i, 0 \leq r \leq n$, le nombre de succès observés : la loi $B(n, \pi : r)$, qui est l'expansion du binôme unitaire $(\pi + \omega)^n$, fournit $p(r)$, la probabilité d'observer r succès, où :

$$p(r) = \binom{n}{r} \pi^r \omega^{n-r}, \quad r = 0, 1, \dots, n. \quad (1)$$

La loi de succession des probabilités $p(r)$ est :

$$p(r) = p(r-1) \frac{\pi}{\omega} \cdot \frac{n+1-r}{r}, \quad (2)$$

à partir de $p(0) = \omega^n$.

L'exemple A, ci-dessus, présente $n = 9$ événements marqués de $r = 5$ réussites. Posant le paramètre π ($0 < \pi < 1$) et $\omega = 1 - \pi$, la probabilité d'un tel ensemble de résultats est, selon (1), $p(r = 5) = {}_9C_5 \cdot \pi^5 \omega^{9-5}$. Supposant qu'il s'agisse d'un jeu de Pile ou Face réalisé avec un jeton honnête, nous aurions $\pi = \omega = 1/2$, et $p = {}_9C_5 \cdot (1/2)^9 \approx 0,2461$.

La moyenne (μ') et les autres moments de r sont :

$$\mu' = n \cdot \pi \quad (3a)$$

$$\sigma^2 = n \cdot \pi \cdot \omega \quad (3b)$$

$$\gamma_1 = (\omega - \pi) / \sqrt{n \cdot \pi \cdot \omega} \quad (3c)$$

$$\gamma_2 = (1 - 6\pi\omega) / (n \cdot \pi \cdot \omega) \quad (3d)$$

Estimation de π . L'estimateur usuel de π pour la loi binomiale est :

$$\hat{\pi} = r / n, \quad (4)$$

soit la proportion de succès observés.

Blaise Pascal (1623-1662) avait « redécouvert » cette loi de distribution pour un jeton honnête, soit le cas particulier avec $\pi = \omega = 1/2$, une loi déjà connue, semble-t-il, depuis l'Antiquité et notamment en Chine. C'est Jacques Bernoulli (1654-1705) qui a trouvé la loi générale et en a formulé la théorie. Et c'est Abraham Moivre (1667-1754) qui, en cherchant à approximer la probabilité $p(r : n, \pi)$ pour n croissant, a le premier trouvé la forme de la loi normale.

La loi de Pascal, la loi géométrique. Maintenant, si, au lieu de fixer le nombre d'essais réalisés, on fixe plutôt le nombre de succès requis, on obtient une autre loi, soit la loi de Pascal, $Pa(r, \pi : n)$, qui énumère la probabilité d'obtenir les r succès voulus au n^e essai, où $n = r, r+1, r+2, \dots$. Dans le cas où un seul succès est requis, en fait le premier succès, nous avons la loi géométrique, $G(\pi : n) = Pa(1, \pi : n)$. La probabilité d'observer les premiers r succès au n^e essai est fournie par (Johnson, Kotz et Kemp, 1992) :

$$p(n) = \binom{n-1}{r-1} \pi^r \omega^{n-r}, \quad n = r, r+1, r+2, \dots \quad (5)$$

La loi de succession des probabilités $p(n)$ est :

$$p(n) = p(n-1) \frac{n}{n+1-r} \omega, \quad (6)$$

à partir de $p(n=r) = \pi^r$.

Reprenons l'Exemple A ci-dessus, et posons $r = 3$: nous sommes intéressé au numéro d'essai auquel nous obtiendrons les 3 premiers succès. Nous voyons ici que le 3^e succès advient à l'essai 6, d'où (pour cette loi) $n = 6$, et la probabilité d'une telle occurrence est $p(n=6) = {}_6C_{3-1} \pi^3 \omega^{6-3}$. Dans le cas d'un jeu de Pile ou Face à jeton honnête, soit $\pi = \omega = 1/2$, cette probabilité s'évaluera comme ${}_5C_2 (1/2)^6 \approx 0,1563$.

Les moments de n , le nombre d'essais exigés, sont :

$$\mu' = r / \pi \quad (7a)$$

$$\sigma^2 = r \cdot \omega / \pi^2 \quad (7b)$$

$$\gamma_1 = (2 - \pi) / \sqrt{r \cdot \omega} \quad (7c)$$

$$\gamma_2 = r^{-1} [6 + \pi^2 / \omega] \quad (7d)$$

Estimation de π . L'estimateur usuel de π pour la loi de Pascal est :

$$\hat{\pi} = (r-1) / (n-1). \quad (8)$$

Ainsi que l'indique le 3^e moment γ_1 , cette distribution a une asymétrie positive prononcée. Le lecteur sera peut-être intéressé de savoir que la loi de Pascal (pour tout r) et la loi géométrique (pour $r = 1$) sont, en temps discret, les équivalents respectifs de la loi Gamma et de la loi exponentielle en temps continu.

La littérature inscrit la loi de Pascal sous le vocable de loi Binomiale négative, laquelle admet un argument r réel (plutôt que seulement entier) : voir Johnson, Kotz et Kemp (1992)

Les lois de Pascal restreintes. Notre étude concerne un processus de Bernoulli dont lequel le nombre de succès à obtenir (r) est prédéterminé, comme dans la loi de Pascal, mais sous la restriction que ces succès apparaissent dans un horizon borné, soit en dedans de k essais, où $k \geq r$. On veut donc étudier le numéro d'essai n auquel r succès seront apparus dans les k plus récents essais, ce pour la première fois : c'est la loi de Pascal restreinte, qu'on peut noter $Pa(r, k, \pi : n)$.

Un processus de Bernoulli tel que celui décrit par la loi de Pascal restreinte se retrouve dans certaines situations de test, dans lesquelles le sujet (ou le mécanisme) testé doit « réussir » r fois en dedans de k essais, le numéro d'essai atteint étant une mesure de sa capacité. On peut concevoir aussi un système physique pour le fonctionnement ou le déclenchement duquel les essais « réussis » doivent être groupés, comme c'est le cas de la sommation spatio-temporelle des influx synaptiques dans la transmission nerveuse (voir aussi Koutras, 1996).

L'étude de la loi de Pascal restreinte n'en est qu'à ses débuts, et seuls quelques cas ou aspects particuliers en ont été vraiment investigués. Faisons exception des cas particuliers déjà connus, soit la loi $Pa(r, \infty, \pi : n)$, soit la loi à horizon k illimité identique à la loi de Pascal ainsi que la loi géométrique $G(\pi : n) = Pa(1, *, \pi : n)$, pour laquelle l'horizon n'importe pas. Les cas spécifiques sont :

- la loi $Pa(r, r, \pi : n)$, aussi appelée loi des succès consécutifs, pour laquelle de beaux résultats sont disponibles ;
- la loi $Pa(2, k, \pi : n)$, soit la loi restreinte appliquée à un groupe de $r = 2$ succès inscrits dans un horizon de $k \geq 2$ essais, cette loi étant aussi bien établie ;
- la loi $Pa(r, k, \pi : n)$, où $k > r > 2$, soit le cas général de la loi de Pascal restreinte, et pour laquelle n'existent en fait que des résultats partiels et une belle solution numérique par la modélisation en chaîne de Markov.

Nous examinerons chacune des lois susnommées à tour de rôle.

La loi des succès consécutifs, $Pa(r, r, \pi : n)$

À quel numéro d'essai observerons-nous r succès consécutifs pour la première fois? C'est le problème, numéroté LXXXVII, auquel s'est attaqué Abraham Moivre dans son fameux traité « The doctrine of chances », en 1738 : « To find the probability of throwing a chance assigned a given number of times without intermission, in any given number of trials » : Moivre donne de ce problème une

solution par développement en série. Le problème a été repris par Uspensky (1937), qui en obtient une fonction génératrice, une approximation reprise par Feller (1968) et une équation de récurrence pour la fonction de répartition. La théorie mathématique des suites présentée par Mood (1940) offre quant à elle un traitement par analyse combinatoire et permet de déterminer la fonction de masse, traitement que nous reproduisons sommairement ici. Quant aux moments, nous les établirons par la méthode de conditionnement suggérée par Ross (1980). Plus récemment, Koutras (1995, 1996, voir aussi Balakrishnan et Koutras, 2002) a étudié cette loi dans un contexte élargi, celui des sommes mobiles (« scans ») en cours de processus, lui reconnaissant le nom de « loi géométrique d'ordre k ». Grant (1946), Mosteller (1941) et Bradley (1968) évoquent des applications spécifiques pour cette loi.

Pour simplifier la présentation dans cette section, nous noterons par $p(n)$ et $P(n)$ respectivement la fonction de masse et la fonction de répartition de la loi $Pa(r, r, \pi : n)$. L'exemple suivant illustre le processus, avec $r = 3$ essais consécutifs requis :

(Exemple B)

Événement	1	2	3	4	5	6	7	8	9	10	11	...
Résultat	1	0	1	0	0	1	1	0	1	1	1	...

le couronnement, c.-à-d. la première suite de $r = 3$ succès, s'observant à l'essai $n = 11$.

La fonction de masse, $p(n)$. L'obtention de r succès consécutifs, – ce qu'on pourra appeler le « couronnement » de l'expérience –, est marquée par trois grandes phases, selon le numéro d'essai n :

- $n < r$: le nombre d'essais est insuffisant, d'où $p(1) = p(2) = \dots = p(r-1) = 0$;
- $n = r$: les r premiers essais sont réussis, d'où $p(r) = \pi^r$;
- $n > r$: il y a 1 ou quelques échecs précédant la suite finale, d'où $p(n) = (1 - \Sigma) \omega \pi^r$.

Dans la troisième phase, notée symboliquement $(1 - \Sigma) \omega \pi^r$, le signe de sommation (Σ) dénote et récapitule les occasions de couronnement antérieures, c.-à-d. la somme des probabilités d'observer une ou quelques suites d'au moins r succès consécutifs avant les $r+1$ essais qui couronnent et terminent la série. Cette phase peut être décomposée à son tour. En effet, écartant les $r+1$ derniers essais, il reste $n-r-1$ essais à considérer. Si ce nombre, $n-r-1$, est inférieur à r , alors la probabilité (Σ) d'un couronnement antérieur est nulle : cette sous-phase a donc une probabilité constante, égale à $\omega \pi^r$. Si le nombre $n-r-1$ atteint r mais n'excède pas $2r$, l'apparition de 1 suite de r succès ou plus est possible, d'où la probabilité (Σ) d'un arrêt antérieur n'est pas nulle. Enfin, si $n-r-1$ déborde $2r$, la probabilité globalisée dans Σ doit tenir compte de l'occurrence possible de 2 ou plus de 2 suites d'au moins r succès. L'analyse

suiuante, inspirée de Mood (1940), est clarifiée dans Bradley (1968, p. 255-257).

Soit $n = t + r + 1$, un numéro d'essai, où $t = n - r - 1$ représente les essais précédant la suite finale : nous cherchons à déterminer la probabilité qu'il se soit produit une ou plusieurs suites d'au moins r succès dans les t premiers essais. Supposons en général i succès et $t - i$ échecs, la somme Σ cherchée correspond à :

$$\Sigma[t, r] = \sum_{i=0}^t c_{i,r} \pi^i \omega^{t-i}, \quad (9)$$

où $c_{i,r}$ indique le nombre d'arrangements contenant au moins une suite d'au moins r succès.

Considérant des suites de r succès, on peut placer k suites de ${}_{t-i+1}C_k$ façons parmi les $t - i + 1$ partitions délimitées par les $t - i$ échecs. Dans chaque cas, il reste $i - k \cdot r$ succès à mêler avec les $t - i$ échecs, ce qui se fait de ${}_{t-k \cdot r}C_{t-i}$ façons, d'où le nombre d'arrangements contenant k suites d'au moins r succès est :

$$A_{i,k} = \binom{t-i+1}{k} \binom{t-k \cdot r}{t-i}. \quad (10)$$

Dans cette formule, les k suites considérées consomment $u \geq k \cdot r$ succès, les $i - u$ succès restants pouvant eux aussi former d'autres suites.

Donnons deux exemples de cette situation, en tablant sur des suites d'au moins $r = 4$ succès.

Dans le premier exemple, posons $n = 13$ essais, comme : 10110110 01111, la série étant fermée par la séquence terminale 01111 indiquée en italique. Il reste donc $t = n - r - 1 = 8$ essais à réarranger, lesquels comptent en général i succès et $t - i$ échecs. Si $i < r$, alors aucun réarrangement ne peut contenir la suite cherchée (ici, 1111 pour $r = 4$). Si $i \geq r$, alors, outre la suite de r succès, il reste $i - r$ succès à mêler aux $t - i$ échecs, ce qu'on peut faire de ${}_{t-i}C_{i-r}$ façons, p. ex. 11110001 01111. De plus, puisque $t - i \leq t - r = 4 \leq r$, alors seule une telle suite est possible (puisque, même avec $t = 2r = 8$ succès, il faudrait un échec intercalé pour former une seconde suite distincte). Cette suite, qu'il faut concevoir comme un bloc, peut s'inscrire avant, parmi ou après les $t - i$ échecs, ici 3, donc en $t - i + 1$ positions : l'exemple fourni admet donc une séquence construite de ${}_{t-i}C_{i-r} = {}_4C_3 = 4$ façons et occupant l'une de $t - i + 1 = 8 - 4$ positions, soit $4 \times 4 = 16$ arrangements différents possibles.

Dans le second exemple, posons $n = 20$ essais, comme : 111011101110111 01111, d'où $t = 15$ essais à réarranger. Si ces t essais comportent $i \geq k \cdot r$ ($k = 1, 2, \dots$) succès, les arrangements pourront présenter de 1 à k suites d'au moins r succès, lesquelles pourront être insérées parmi les $t - i$ échecs restants, ce de ${}_{t-i+1}C_k$ manières ; or, il y a ${}_{t-k \cdot r}C_{t-i}$ arrangements différents des succès et échecs produisant au moins k suites, d'où le nombre total d'arrangements est donné par le produit (10). Dans notre exemple, $i = 12 \geq 3 \cdot r =$

12, d'où de 1 jusqu'à $k = 3$ suites sont possibles, comme 111101111011110, celles-ci pouvant être arrangées de ${}_{3+1}C_3 = 4$ façons, soit 0S0S0S, S0S0S0, S0S0S et S0S00S (où S \equiv 1111).

La formule (10) ci-dessus compte les arrangements distincts comportant au moins k suites d'au moins r sans considérer la ou les suites surnuméraires que les succès résiduels (ceux non comptabilisés dans les k suites calculées) peuvent engendrer. Or, notre somme de probabilités (9) doit compter toutes les suites possibles, c.-à-d. tous les arrangements distincts des t essais contenant exactement k suites d'au moins r essais. Soit $B_{i,k}$ ce nombre. Prenons l'exemple de $k = 3$ suites. Ni $B_{i,1}$ ni $B_{i,2}$ ne participent aux arrangements décrits par $A_{i,3}$ (lesquels contiennent au moins $k = 3$ suites). D'autre part, $B_{i,3}$ y participe d'une façon, les 3 suites dans $B_{i,3}$ étant comptées telles quelles dans $A_{i,3}$. Maintenant, $B_{i,4}$ participe à $A_{i,3}$ de ${}_{4}C_3 = 4$ façons, selon les combinaisons de suites 1-2-3, 1-2-4, 1-3-4 et 2-3-4, et généralement $B_{i,j}$ y participe de ${}_jC_3$ façons. Nous avons donc la relation générale :

$$A_{i,k} = B_{i,k} + \binom{k+1}{k} B_{i,k+1} + \binom{k+2}{k} B_{i,k+2} + \dots, \quad (11)$$

somme qui représente le nombre d'arrangements contenant 1, 2, 3, ... suites d'au moins r succès. Or, c'est le coefficient c_i dans (9) qui nous intéresse afin de déterminer Σ , et il est donné par l'addition simple des composantes $B_{i,k}$, soit $c_i = B_{i,1} + B_{i,2} + \dots$. Disposant les expressions (11) en tableau,

$$\begin{aligned} A_{i,1} &= B_{i,1} + 2B_{i,2} + 3B_{i,3} + 4B_{i,4} + \dots \\ A_{i,2} &= B_{i,2} + 3B_{i,3} + 6B_{i,4} + \dots \\ A_{i,3} &= B_{i,3} + 4B_{i,4} + \dots \\ A_{i,4} &= B_{i,4} + \dots \end{aligned}$$

nous voyons que la somme voulue est :

$$c_{i,r} = A_{i,1} - A_{i,2} + A_{i,3} - A_{i,4} + \dots \quad (12)$$

Pour notre exemple B, où nous avons $i = 12$ succès et $t - i = 3$ échecs, $A_{12,1} = 165$, $A_{12,2} = 35$ et $A_{12,3} = 4$, d'où $c_{12,4} = 165 - 35 + 4 = 134$, la probabilité associée à cet ensemble de succès et d'échecs étant $134 \times \pi^{12} \times \omega^3$.

L'ensemble des résultats précédents permet d'établir l'expression suivante pour le calcul de la probabilité $p(n,r)$ d'obtenir au moins une suite d'au moins r succès une première fois à l'essai n :

$$p(n,r) = \begin{cases} 0 & \text{si } n < r, \\ \pi^r & \text{si } n = r, \\ \omega \cdot \pi^r & \text{si } r < n < 2r, \end{cases} \quad (13)$$

$$\left[1 - \sum_{i=2r}^{n-1} \pi^{i-r} \omega^{n-1-i} \times \right.$$

$$\left. \left(\sum_{j=1}^{\lfloor (i-r)/r \rfloor} (-1)^{j-1} \binom{n-i}{j} \binom{n-r-1-j \cdot r}{n-1-i} \right) \right] \omega \pi^r$$

si $2r < n$.

La fonction de répartition, $P(n)$. Considérons maintenant le

calcul de la fonction de répartition $P(n)$, équivalant à $\sum_{u=r}^n p(u)$. Les valeurs initiales de la distribution sont évidemment $P(1) = P(2) = \dots = P(r-1) = 0$ et $P(r) = \pi^r$. Pour $n > r$, la probabilité que le processus termine à l'essai $n+1$ suppose qu'il n'ait pas abouti avant l'essai $n - r$ et qu'un échec survienne alors, suivi d'une suite de r succès, d'où la probabilité conjointe est :

$$p(n+1) = [1 - P(n-r)] \omega \pi^r, \quad (14)$$

la fonction de répartition étant alors obtenue récursivement, par :

$$P(n+1) = P(n) + [1 - P(n-r)] \omega \pi^r, \quad (15)$$

les probabilités individuelles découlant alors des différences successives de (15), soit :

$$p(n, r) = P(n) - P(n-1). \quad (16)$$

Uspensky (1937), Feller (1968) et Laurencelle (1987) proposent des approximations pour ces fonctions, ce dernier au moyen d'une loi géométrique.

Les moments. Laurencelle (1987) utilise une méthode de conditionnement pour déterminer les moments de cette loi. Illustrons cette méthode par une simple loi géométrique, $G(\pi : n)$, associée au tout premier succès. On peut obtenir ce succès dès l'essai $n = 1$, avec probabilité π , ou sinon, avec probabilité ω , le processus est prolongé d'un essai, soit $1 + n$, d'où le s^{e} moment de n est :

$$E(n^s) = \pi \cdot 1^s + \omega \cdot E(1+n)^s \quad (17)$$

Par exemple, pour $s = 1$, nous obtenons $E(n) = \pi \cdot 1 + \omega \cdot [1 + E(n)]$, d'où $E(n) = 1/\pi$, tandis que $E(n^2) = \pi \cdot 1^2 + \omega \cdot [1 + 2E(n) + E(n^2)] = (2 - \pi) / \pi^2$, et $\sigma^2(n) = \omega / \pi^2$, etc. Les cas un peu plus complexes peuvent s'exprimer aussi sous forme d'une arborescence : voir p. ex. Laurencelle (1983).

Pour la loi des r succès consécutifs, nous conditionnons sur le premier échec, puis posons :

$$E(n^s) = \pi^r \cdot r^s + \omega \cdot E(1+n)^s + \omega \cdot \pi \cdot E(2+n)^s + \dots + \omega \cdot \pi^{r-1} \cdot E(r+n)^s. \quad (18)$$

Isolant le terme $E(n^s)$ et simplifiant, nous obtenons les moments :

$$\mu' = E(n) = \frac{1 - \pi^r}{\omega \cdot \pi^r} \quad (19a)$$

$$\sigma^2 = \frac{1 - \pi^{2r+1} - (2r+1) \cdot \omega \cdot \pi^r}{(\omega \cdot \pi^r)^2} \quad (19b)$$

$$\gamma_1 = \frac{2 - 3(2r+1)\rho + (1+3r+3r^2)\rho^2 - \pi^{3r+1}(1+\pi)}{[1 - (2r+1)\rho - \pi^{2r+1}]^{3/2}} \quad (19c)$$

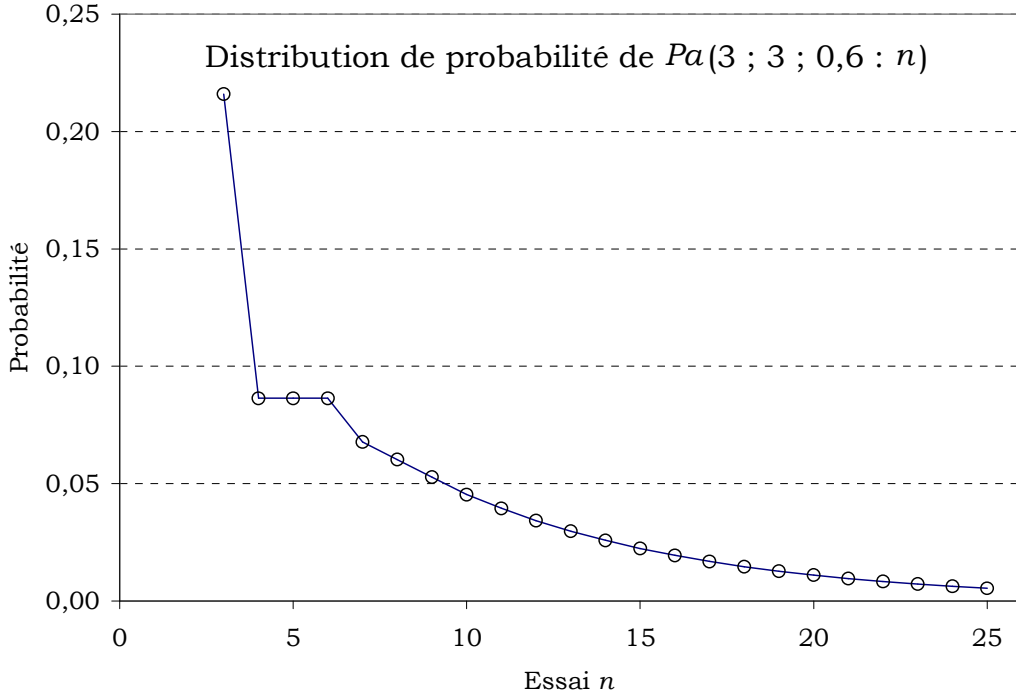


Figure 1. Distribution de probabilité $p(n)$ de la loi $Pa(3 ; 3 ; 0,6 : n)$.

$$\gamma_2 = \frac{6 - 12(2r + 1)\rho + (7 + 24r + 24r^2)\rho^2 - (1 + 4r + 6r^2 + 4r^3)\rho^3 - \pi^{4r+1}(1 + 4\pi + \pi^2)}{(1 - (2r + 1)\rho - \pi^{2r+1})^2} \quad (19d)$$

où $\rho = \omega \cdot \pi^r$ (dans les deux dernières expressions). On montre facilement que l'espérance (μ') de cette loi est plus forte que celle de la loi de Pascal, tout comme pour la variance.

En guise d'illustration, prenons une loi $Pa(3 ; 3 ; 0,6 : n)$, soit une loi à $r = k = 3$ succès consécutifs, chaque succès ayant probabilité $\pi = 0,6$. La distribution de probabilité est illustrée à la figure 1, ci-dessus. Forte asymétrie positive, bonne concentration des données dans les petites valeurs de n , grande dispersion : ces constatations se répercutent dans les moments de la variable, soit :

$$\mu' \approx 9,0741 ; \sigma^2 \approx 49,1907 ; \gamma_1 \approx 2,0319 ; \gamma_2 \approx 6,1316 .$$

Quant aux probabilités individuelles, celle d'un succès immédiat, dès le 3^e essai, est $p(3) = \pi^3 = 0,6^3 = 0,216$. Pour les $r = 3$ essais subséquents, $n = 4, 5$ et 6 , $p(n) = \omega \cdot \pi^3 = 0,0864$, avec $\omega = 1 - \pi = 0,4$. Illustrons la partie plus complexe du calcul, avec $n = 12$. De ces 12 essais, on écarte d'abord les $r + 1 = 4$ essais terminaux, soit « 0111 », de probabilité conjointe $\omega \cdot \pi^3$. Il reste $t = 8$ essais à considérer, qui auraient pu occasionner des suites antérieures de $r = 3$ succès ou plus (comptabilisées dans la somme Σ) et qu'il faut défalquer. Ces 8 essais peuvent contenir de $u = 0$ à 8 réussites, et la contribution de chaque ensemble doit être comptée. En voici les valeurs :

$$p(0) = p(1) = p(2) = 0 ; p(3) = 0,013271 ;$$

$$p(4) = 0,082944 ; p(5) = 0,119066 ; p(6) = 0,201554 ;$$

$$p(7) = 0,089580 ; p(8) = 0,016796 .$$

Pour les essais comportant $u = 0$ à 2 succès, aucune suite ne peut en être formée, d'où $p(u) = 0$ pour ces cas. Pour $u = 3$, il y a 1 suite (de 3) possible, à placer parmi les $t - u = 5$ échecs, d'où $t - u + 1 = 6$ arrangements, et une probabilité de $6 \cdot \pi^3 \omega^5 = 0,013271$. Prenons $u = 7$: ici, on peut former 1 suite ou 2 suites. Pour 1 suite, le facteur $A_{u,k} = A_{7,1} = {}_{8-7+1}C_1 \times {}_{8-1 \times 3}C_{8-7} = {}_2C_1 \times {}_5C_1 = 2 \times 5 = 10$; pour 2 suites, ${}_{8-7+1}C_2 \times {}_{8-2 \times 3}C_{8-7} = {}_2C_2 \times {}_2C_1 = 1 \times 2 = 2$; le coefficient $c_{7,3} = 10 - 2 = 8$, et la probabilité cherchée, $8 \cdot \pi^7 \cdot \omega^1 = 0,089580$. La somme (Σ) des 8 probabilités ci-dessus étant 0,603211, la probabilité finale d'obtenir pour la première fois la suite « 111 » au 12^e essai est $(1 - 0,603211) \times \omega \cdot \pi^3 \approx 0,0342826$.

La fonction de répartition par récursion (15) permet d'obtenir plus facilement les mêmes résultats, en appliquant la différence (16) entre les valeurs successives.

Estimation de π . Il ne semble pas y avoir d'estimateur connu de π pour la loi des succès consécutifs : l'utilisateur est alors contraint de trouver (par repérage) l'estimation du maximum de vraisemblance en développant explicitement la fonction de masse (13) ou encore en inversant numériquement l'expression d'espérance (19a) : voir la

Tableau 1. Coefficients $C_x(n)$ indiquant le nombre de chaînes couronnées d'un processus $Pa(2,k)$ comportant x succès, pour les numéros d'essais $n = 2$ à $n = 12$

$n \setminus x$	$Pa(2,2)$						$Pa(2,3)$				$Pa(2,4)$			$Pa(2,5)$		
	2	3	4	5	6	7	2	3	4	5	2	3	4	2	3	4
2	1						1				1			1		
3	1						2				2			2		
4	1	1					2				3			3		
5	1	2					2	1			3			4		
6	1	3	1				2	3			3	1		4		
7	1	4	3				2	5			3	3		4	1	
8	1	5	6	1			2	7	1		3	6		4	3	
9	1	6	10	4			2	9	4		3	9		4	6	
10	1	7	15	10	1		2	11	9		3	12	1	4	10	
11	1	8	21	20	5		2	13	16	1	3	15	4	4	14	
12	1	9	28	35	15	1	2	15	25	5	3	18	10	4	18	1

section suivante.

La loi de Pascal restreinte pour 2 succès, $Pa(2, k, \pi : n)$

La loi $Pa(2, k, \pi : n)$ concerne le numéro d'essai n auquel, pour la première fois, deux succès sont enregistrés en dedans de k essais, où $k \geq 2$. Posant $k = 2$, nous retrouvons le cas des succès consécutifs traité plus haut. Si $k > 2$, nous faisons face au premier exemple de la loi de Pascal restreinte générale, qui est aussi le seul exemple (presque) complètement résolu. L'article de Koutras (1996) en donne un traitement magistral ; d'autres données apparaissent aussi dans Laurencelle (1998a). Supposons un processus $Pa(2,4)$, dans lequel nous recherchons 2 succès inscrits dans un intervalle de 4 essais. Dans l'exemple C, ci-dessous, nous voyons que le processus $Pa(2,4)$ se conclut à l'essai 11.

(Exemple C)

Événement	1	2	3	4	5	6	7	8	9	10	11	...
Résultat	1	0	0	0	1	0	0	0	1	0	1	...

La fonction de masse, $p(n)$. La distribution des probabilités $p(n)$ peut s'articuler simplement autour des coefficients $C_x(n)$ dans l'expression :

$$p(n) = C_2(n) \cdot \omega^{n-2}\pi^2 + C_3(n) \cdot \omega^{n-3}\pi^3 + \dots \quad (20)$$

Les premières valeurs de ces coefficients pour des processus à $k = 2$ à 5 sont présentées au tableau 1.

L'examen des ensembles de coefficients qui apparaissent dans le tableau 1 révèle des régularités qui correspondent aux formules générales suivantes :

$$C_2 = k - 1 \quad (21)$$

$$C_3 = \binom{n - (k - 1) - 1}{2} - \binom{n - 2(k - 1) - 1}{2}$$

$$C_4 = \binom{n - 2(k - 1) - 1}{3} - \binom{n - 3(k - 1) - 1}{3}$$

$$C_4 = \binom{n - 3(k - 1) - 1}{4} - \binom{n - 4(k - 1) - 1}{4}$$

etc.,

lesquelles permettent de générer la distribution complète.

Il est aussi possible de façonner une fonction récurrente, basée sur la fonction cumulative et amorcée à partir des k premiers essais. Dès que $n \geq k$, le processus $Pa(2, k)$ présente $k-1$ séquences terminales, soit les séquences $T_1 = 0^{k-2}11$, $T_2 = 0^{k-3}101$, $T_3 = 0^{k-4}10^21..$, $T_{k-1} = 10^{k-2}1$, chacune de probabilité $\omega^{k-2}\pi^2$. Pour qu'une chaîne soit proprement couronnée à l'essai n , elle doit débiter par une sous-chaîne non couronnée de longueur $n-k-i$ essais, suivie d'un échec puis de la séquence terminale T_i . La probabilité qu'une chaîne de $n-k-i$ essais soit non couronnée est, par définition, $1 - P(n-k-i)$, probabilité qu'on doit joindre à celle, égale à ω , de l'échec intercalé et à celle, égale à $\omega^{k-2}\pi^2$, de la séquence T_i . Additionnant les chaînes qui confluent à l'essai n , nous obtenons la fonction récurrente :

$$p(n) = [1 - P(n-k-1)] \cdot \omega^{k-1}\pi^2 \quad (22)$$

$$+ [1 - P(n-k-2)] \cdot \omega^k\pi^2$$

$$+ \dots$$

$$+ [1 - P(n-2k+1)] \cdot \omega^{2k-3}\pi^2,$$

valide dès $n \geq 2k-1$. En complément, pour les numéros n s'étendant de 2 à k , nous avons $p(n) = (n-1) \cdot \omega^{n-2}\pi^2$ et, pour $k+1 \leq n < 2k-1$, $p(n) = (k-1) \cdot \omega^{n-2}\pi^2 + \frac{1}{2}(n-k)(n-k-1) \cdot \omega^{n-3}\pi^3$.

Koutras (1996) propose une fonction récurrente plus simple, soit :

$$p(n) = \omega \cdot p(n-1) + \pi \cdot \omega^{k-1} p(n-k), \quad (23)$$

avec, comme tantôt, $p(n) = (n-1) \cdot \omega^{n-2} \pi^2$ pour $2 \leq n \leq k$.

La fonction récurrente (23) ci-dessus donne lieu à la fonction génératrice de probabilité :

$$\phi(z) = \frac{(\pi \cdot z)^2 [1 - \omega \cdot z]^{k-1}}{(1 - \omega \cdot z - \pi \cdot \omega^{k-1} z^k)(1 - \omega \cdot z)}, \quad |z| \leq 1, \quad (24)$$

dont les dérivations successives fournissent les moments de la variable.

La fonction de répartition $P(n)$ s'obtient en cumulant des probabilités simples, $p(n)$.

Les moments. On peut obtenir les moments de la loi $Pa(2, k)$ par conditionnement. Prenons l'exemple d'un processus $Pa(2,3)$, donnant lieu à l'arborescence :

$$n^s = n^s [\pi + \omega]^{A(2,3)} \quad (25)$$

soit, ici :

$$n = [\pi^2](2^s) + [\omega\pi^2](3^s) + [\omega](1+n)^s + [\omega^2\pi](3+n)^s. \quad (26)$$

L'expansion, puis la réduction de (26) fournit les moments à l'origine μ_s' puis les indices classiques. Cette méthode ainsi que les dérivées de la fonction $\phi(z)$ évaluées à $z = 1$ fournissent ces indices :

$$\mu_1' = \frac{2 - \omega^{k-1}}{\pi(1 - \omega^{k-1})} \quad (27a)$$

$$\sigma^2 = \frac{\omega}{\pi^2} \left[1 + \frac{(2k-1)\pi \cdot \omega^{k-2}}{(1 - \omega^{k-1})^2} \right]. \quad (27b)$$

Quant aux moments supérieurs, plus complexes, on peut les obtenir par une récurrence élaborée par Koutras (1996) pour les moments à l'origine, soit, en posant $\mu_0' = 1$:

$$\mu_s' = \frac{1}{\pi(1 - \omega^{k-1})} \times \left\{ \pi^2 \sum_{j=2}^k j^s \omega^{j-2} + \omega \sum_{i=1}^{s-1} \binom{s}{i} (1 + \pi \cdot \omega^{k-1} k^{s-i}) \mu_i' \right\}. \quad (28)$$

$s > 1$.

Koutras (1996) place le mode de la loi $Pa(2, k)$ à l'essai $1 + \lfloor \min(k-1, \pi^{-1}) \rfloor$.

La figure 2 donne l'allure de la distribution des probabilités d'un processus $Pa(2; 4; \pi = 0,6)$. Forte asymétrie positive et concentration de probabilité aux basses valeurs, les moments de cette distribution sont :

$$\mu_1' \approx 3,447293 ;$$

$$\sigma^2 \approx 3,231630 ;$$

$$\gamma_1 \approx 2,185444 ;$$

$$\gamma_2 \approx 7,162679 .$$

Pour obtenir, par exemple, γ_1 , nous devons d'abord avoir les moments $\mu_1' \approx 3,447293$ et $\mu_2' = (\mu_1')^2 + \sigma^2 \approx 15,115459$. Puis, utilisant (28) : $\mu_3' = [0,6 \times (1 - 0,4^3)]^{-1} \{ 0,6^2 \cdot (2^3 \times 0,4^0 + 3^3 \times 0,4^1 + 4^3 \times 0,4^2) + 0,4 \times [{}_3C_0 \times (1 + 0,6 \times 0,4^2 \times 4^3) \cdot 1 + {}_3C_1 \times (1 + 0,6 \times 0,4^2 \times 4^2) \cdot 3,447293 + {}_3C_2 \times (1 + 0,6 \times 0,4^2 \times 4^1) \cdot 15,115459] \} \approx 0,5616^{-1} \times [0,36 \times 29,04 + 0,4 \times 96,130391] \approx 87,084324$. De là, le 3^e moment centré est $\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1')^3 \approx 12,696157$, puis l'indice d'asymétrie $\gamma_1 = \mu_3/\sigma^3 \approx 12,696157 / 3,231630^{3/2} \approx 2,185448$.

Quant à la distribution de probabilité, les premières valeurs sont $p(0) = p(1) = 0$; $p(2) = \pi^2 = 0,6^2 = 0,36$; $p(3) = 2\omega\pi^2 \approx 0,288$; $p(4) = 3\omega 2\pi^2 \approx 0,1728$. Pour la suite, soit $n > k = 4$, la série (23) récurrente de Koutras donne, par exemple :

$p(5) = \omega \cdot p(4) + \pi \cdot \omega^3 \cdot p(1) = 0,4 \times 0,1728 + 0,6 \times 0,4^3 \times 0 = 0,06912$
ou

$$\begin{aligned} p(10) &= \omega \cdot p(9) + \pi \cdot \omega^3 \cdot p(6) \\ &\approx 0,4 \times 0,009732 + 0,6 \times 0,4^3 \times 0,041472 \\ &\approx 0,005485. \end{aligned}$$

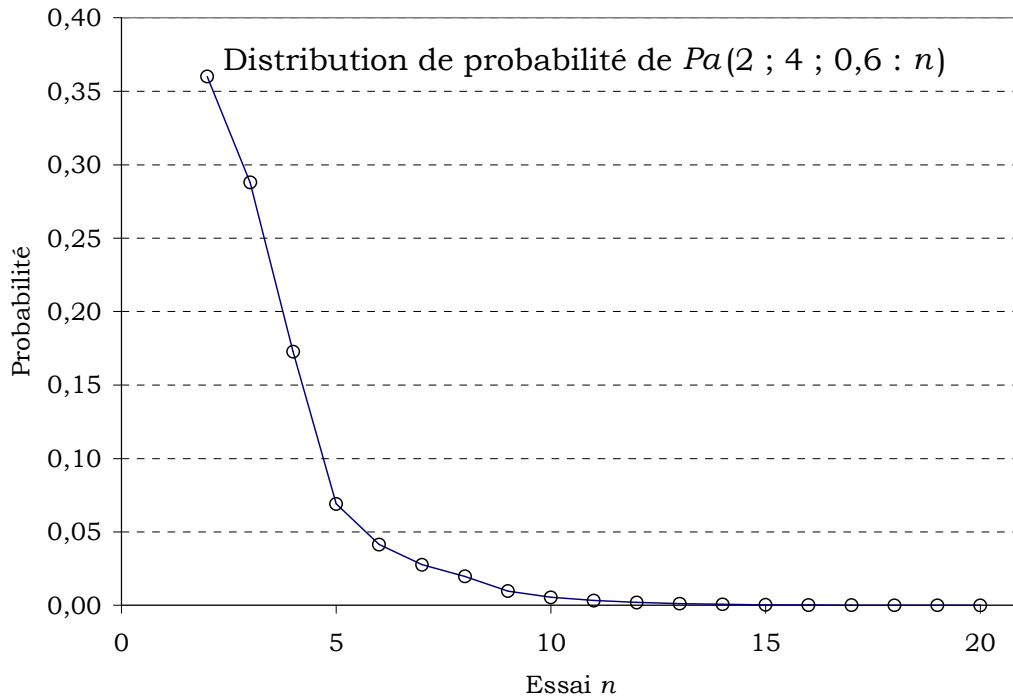


Figure 2. Distribution de probabilité $p(n)$ de la loi $Pa(2 ; 4 ; 0,6 : n)$.

Estimation de π . À défaut d'estimateurs appropriés du paramètre π pour la loi de Pascal, $Pa(r, k, \pi : n)$, il reste deux procédés universels disponibles pour quantifier ce paramètre : l'inversion du moment 1 (μ_1') lorsque celui est disponible, et l'estimation directe du maximum de vraisemblance, $p(n_1) \cdot p(n_2) \cdot \dots \cdot p(n_m)$ à partir de m réalisations (n_i) du processus. Pour des fins d'illustration, supposons encore un processus $Pa(2,4)$ et les réalisations suivantes :

$$n_1 = 4 ; n_2 = 8 ; n_i = (4, 8, 5, 5, 6 : \bar{n}=5,6)$$

Par inversion du moment μ_1' (27a), il s'agit de trouver la valeur de π telle que $(2 - \omega^3) / [\pi(1 - \omega^3)]$ soit égale à n . Par itération, on trouve :

$$\hat{\pi}(4) = 0,529 ; \hat{\pi}(8) = 0,311 ; \hat{\pi}(5,6) = 0,405,$$

l'estimation basée sur la moyenne de 5 valeurs ($\bar{n} = 5,6$) étant évidemment plus sûre.

Avant de calculer directement la vraisemblance (c.-à-d. la probabilité a posteriori) de la réalisation n et en chercher le maximum, il faut d'abord identifier les expressions de probabilité requises par le processus $Pa(2,4)$. Appliquant l'une ou l'autre des fonctions (22) ou (23), voire les coefficients du tableau 1, nous obtenons :

$$\begin{aligned} p(4) &= 3\omega^2\pi^2 ; p(5) = 3\omega^3\pi^2 ; \\ p(6) &= 3\omega^4\pi^2 + \omega^3\pi^3 ; p(8) = 3\omega^6\pi^2 + 6\omega^5\pi^3 . \end{aligned}$$

Ainsi, pour $n_1 = 4$, la valeur de π qui maximise $3\omega^2\pi^2$ est $\hat{\pi} \approx 0,667$, et elle est de 0,309 pour $n_2 = 8$. Quant à l'ensemble de 5 valeurs, la vraisemblance est obtenue par $p(4) \cdot p(5)^2 \cdot$

$p(6) \cdot p(8)$, et son maximum est atteint vers $\pi \approx 0,383$.

Les mêmes méthodes s'appliquent aussi bien à la loi des succès consécutifs comme aux réalisations communes de la loi de Pascal restreinte, pour autant qu'on dispose des expressions algébriques appropriées.

La loi de Pascal restreinte, $Pa(r, k, \pi : n)$: le cas général

Il ne semble pas exister de fonctions connues permettant de générer la distribution de probabilité dans le cas général de la loi de Pascal restreinte. Sont solutionnées les cas particuliers détaillés ci-dessus, soit la loi des succès consécutifs, $Pa(r, r, \pi)$, et la loi restreinte sur 2 succès, $Pa(2, k, \pi)$. Pour les cas où $k > r > 2$, rien n'est connu, sauf quelques premiers moments (Laurencelle, 1998a).

L'évaluation par énumération des chaînes, même à l'aide d'un programme d'ordinateur, se révèle frustrante : les arborescences binaires sont très longues, les moments ne se stabilisent que lentement, et la capacité d'ordinateur est rapidement débordée dès que le paramètre n grimpe, ce qui advient lorsque $r \rightarrow k$ et $\pi \rightarrow 0$. Il reste l'évaluation par la chaîne de Markov.

La représentation de la loi $Pa(r, k, \pi : n)$ par une matrice de Markov. La loi de Pascal restreinte décrit un processus stochastique qui, démarré à l'essai $n = 1$, connaît en fait deux régimes, le premier avant restriction, soit pour les essais $n = r, r+1, \dots, k$, et le second avec restriction, soit pour

les essais $n = k+1, k+2, \dots$. Le calcul de la distribution de probabilité ainsi que celui des moments tiennent compte de cette division. Par exemple, le premier moment, $\mu = \mu_1' = E(n)$, sera constitué par la somme $\mu' + \mu''$, le premier terme représentant la contribution des quelques premiers essais ($r \leq n \leq k$), le second celle des essais ultérieurs ($k < n$). L'application de la chaîne de Markov à ce type de problème est bien documentée (Kemeny et Snell, 1976 ; Balakrishnan et Koutras 2002). Nous présentons les développements issus de Laurencelle (1983).

À partir de l'essai $n = k + 1$, le processus de Bernoulli soumis à la loi de Pascal restreinte subit la restriction, celle d'un horizon des k derniers essais. Soit un essai quelconque de ce régime, l'essai $n+k$, terminant une suite d'essais pour laquelle le critère de r succès n'a pas encore été atteint. Les k derniers essais advenus sont :

$$X_{n+1} \ X_{n+2} \ X_{n+3} \ \dots \ X_{n+k} ,$$

chacun pouvant être un succès ou un échec, jusqu'à concurrence de $r-1$ succès. La réalisation de l'essai suivant, de numéro $n+k+1$, produira donc pour examen la chaîne :

$$X_{n+2} \ X_{n+3} \ \dots \ X_{n+k} \ X_{n+k+1} ,$$

par laquelle le critère est susceptible d'être atteint (si l'essai X_{n+k+1} est un succès et la chaîne comporte $r-1$ succès antérieurs) ou non. Chaque essai ayant une réalisation binaire, d'échec ou de succès, il y a donc 2^k patrons différents de chaînes de type « $X_{n+1} \ X_{n+2} \ X_{n+3} \ \dots \ X_{n+k}$ », qui donnent lieu à 2^k patrons différents de chaînes subséquentes de type « $X_{n+2} \ X_{n+3} \ \dots \ X_{n+k} \ X_{n+k+1}$ », la transition d'une chaîne à la chaîne subséquente dépendant de la valeur de l'essai $n+k+1$: si X_{n+k+1} est un succès, la probabilité de transition est π ; si un échec, elle est $\omega = 1-\pi$. En croisant les 2^k patrons d'entrée et leurs 2^k patrons subséquents et en inscrivant la probabilité de transition, on obtient la matrice de transition, qui représente la progression, nous dirions le roulement, du processus de Bernoulli dans un horizon de k essais.

Pour que le processus étudié par la loi de Pascal restreinte soit représenté spécifiquement par la matrice de probabilités, il faut considérer que, une fois atteint le critère de r succès, la roue s'arrête : certaines probabilités doivent donc être réaménagées, et la matrice doit être rendue « absorbante ».

Soit $c_n = \sum_n^{(k)} x_j$, le nombre de succès observés dans les k plus récents essais, incluant l'essai n , et $c_{n+1} = \sum_{n+1}^{(k)} x_j$, le nombre obtenu à l'essai subséquent $n+1$, où $0 \leq (c_n, c_{n+1}) \leq k$. Il s'agit alors de restructurer la matrice de transition $2^k \times 2^k$ des chaînes $n^{(k)}$ à $n+1^{(k)}$ de façon à obtenir une matrice quadripartite, selon les divisions suivantes :

P : $c_n < r$ et $c_{n+1} < r$	S : $c_n < r$ et $c_{n+1} \geq r$
$c_n \geq r$ et $c_{n+1} < r$	$c_n \geq r$ et $c_{n+1} \geq r$

La première sous-matrice, **P**, inclut les chaînes pour lesquelles $c_n < r$ et $c_{n+1} < r$ (c.-à-d. où le critère n'est pas atteint) et elle correspond au piétinement du processus. La sous-matrice de droite, **S**, avec $c_n < r$ et $c_{n+1} \geq r$, dénote l'atteinte du critère. Une fois le critère atteint, le processus stoppe, de sorte que le roulement dans la quatrième sous-matrice ($c_n \geq r$ et $c_{n+1} \geq r$) est certain, alors qu'il est impossible de se retrouver dans la troisième sous-matrice ($c_n \geq r$ et $c_{n+1} < r$) : afin de dénoter et imposer ces dernières certitude et impossibilité, les sous-matrices sont remplacées respectivement par des matrices identité (**I**) et nulle (**0**), de dimensions appropriées. Nous aboutissons ainsi à la représentation matricielle de la loi $Pa(r, k, \pi)$, soit, symboliquement :

P	S
0	I

Nous fournissons un exemple explicitement développé, plus bas.

Calcul de l'espérance et de la variance. Chaque moment de la variable n , obéissant à la loi de Pascal restreinte $Pa(r, k, \pi)$, est obtenu par l'addition du moment partiel associé au régime initial sans restriction, pour $n < k$, et de l'autre moment partiel, généré par le roulement sous restriction d'un horizon de k essais et représenté par la matrice de Markov $\mathbf{T}(2^k \times 2^k)$. Le moment initial partiel d'ordre t , M_t , relève de la loi de Pascal (5) et s'obtient par :

$$M_t = \sum_{n=r}^k \binom{n-1}{r-1} \omega^{n-r} \pi^r n^t . \tag{29}$$

Soit **V**, un vecteur-colonne de dimension 2^k , qui dénote les probabilités d'entrée dans chaque ligne de la matrice **T** : par exemple, si la ligne 1 représente k échecs, soit $(0_1 \ 0_2 \ \dots \ 0_k)$, nous aurons $\mathbf{V}[1] = \omega^k$. Soit enfin **A** = $(\mathbf{I} - \mathbf{P})^{-1}$. Alors (Laurencelle, 1983),

$$\mu = E(n) = M_1 + k \cdot (1 - M_0) + \alpha \tag{30}$$

et : $\sigma^2 = \text{Var}(n) = M_2 + k^2 \cdot (1 - M_0) + 2k \cdot \alpha + \beta - [E(n)]^2$, (31) où :

$$\alpha = \mathbf{V}^T(\mathbf{A} \cdot \mathbf{1}) ; \beta = \mathbf{V}^T(2\mathbf{A} - \mathbf{I}) \cdot \mathbf{A} \cdot \mathbf{1} . \tag{32}$$

Dans les expressions (32), seule la portion congrue du vecteur **V**, celle correspondant aux lignes de **A** (ou **P**), est utilisée ; \mathbf{V}^T dénote la transposée de **V**.

Calcul des probabilités $p(n)$. Les premiers essais de la loi de Pascal restreinte, ceux situés en-deçà de l'horizon $n = k$, sont gouvernés par la loi de Pascal (5), d'où :

$$p(n) = 0, \ n = 1 \dots r-1 \tag{33}$$

$$= \binom{n-1}{r-1} \omega^{n-r} \pi^r, \ n = r \dots k .$$

À partir de cette borne, soit pour $n = k + 1$ et au-delà, la

Chaîne	0000	0001	0010	0011	0100	0101	0110	1000	1001	1010	1100	0111	1011	1101	1110	1111
0000	ω	π														
0001			ω	π												
0010					ω	π										
0011							ω					π				
0100								ω	π							
0101										ω			π			
0110											ω			π		
1000	ω	π														
1001			ω	π												
1010					ω	π										
1100								ω	π							
0111															ω	π
1011							ω					π				
1101										ω			π			
1110											ω			π		
1111															ω	π

Figure 3. Matrice des probabilités de transition pour des chaînes de $k = 4$ essais de Bernoulli, structurée selon que la chaîne contient moins de $r = 3$ succès ou 3 succès ou plus

Chaîne	0000	0001	0010	0011	0100	0101	0110	1000	1001	1010	1100	0111	1011	1101	1110	1111
0000	ω	π														
0001			ω	π												
0010					ω	π										
0011							ω					π				
0100								ω	π							
0101										ω			π			
0110											ω			π		
1000	ω	π														
1001			ω	π												
1010					ω	π										
1100								ω	π							
0111	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
1011	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
1101	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
1110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
1111	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Figure 4. Matrice absorbante \mathbf{T} du processus de la loi de Pascal restreinte $\text{Pa}(3, 4, \pi)$

« roue » de Markov prend la relève, et nous avons :

$$p(n) = Z(n) - Z(n-1), \quad n = k+1, k+2, \dots \quad (34)$$

$$\text{où : } Z(n) = 0, \quad n = k \quad (35)$$

$$Z(n) = \mathbf{V}^T(\mathbf{S}_{n-k}\mathbf{1}), \quad n = k+1, k+2, \dots$$

La sous-matrice \mathbf{S}_{n-k} dans (35) est obtenue à partir de \mathbf{T}^{n-k} ,

c.-à-d. la $(n - k)^e$ puissance de la matrice absorbante \mathbf{T} .

Illustration des calculs pour $k = 4$ et $r = 3$. La loi $\text{Pa}(4, 3, \pi)$ constitue le premier cas particulier dont le traitement n'est pas accessible autrement que par la chaîne de Markov : nous l'utilisons pour illustrer les calculs.

Tableau 2.

Chaîne	Probabilité d'entrée (V)
0000	ω^4
0001	$\omega^3\pi$
0010	$\omega^3\pi$
0011	$\omega^2\pi^2$
0100	$\omega^3\pi$
0101	$\omega^2\pi^2$
0110	$\omega^2\pi^2$
1000	$\omega^3\pi$
1001	$\omega^2\pi^2$
1010	$\omega^2\pi^2$
1100	$\omega^2\pi^2$
0111	$\omega \cdot \pi^3$
1011	$\omega^2\pi^2$
1101	$\omega \cdot \pi^3$
1110	$\omega \cdot \pi^3$
1111	π^4

Préparation la matrice T. L'horizon $k = 4$ fait apparaître $2^4 = 16$ patrons de chaîne, chaque chaîne contenant 4 essais de Bernoulli. Dénotant chaque essai réalisé par « 1 » pour le succès et par « 0 » pour l'échec, ces patrons sont : 0000, 0001, 0010, 0011, ..., 1110, 1111. Divisons cette liste de patrons selon que le nombre de succès enregistrés atteint ou non le critère de $r = 3$ succès, i.e. selon que $c_n < 3$ ou $c_n \geq 3$. Une fois divisée, la liste permet de définir la condition d'entrée et la condition de succession dans une matrice $2^k \times 2^k$ de Markov, soit celle qui apparaît en tableau à la figure 3.

La matrice se lit ainsi. La suite $(0 \ 0 \ 0 \ 1)_n$ à l'essai n (indiquée à la 2^e ligne de la matrice) devient la suite $(0 \ 0 \ 1 \ x)$ à l'essai $n+1$ (apparaissant à la colonne appropriée). Elle transite donc vers $(0 \ 0 \ 1 \ 0)_{n+1}$ si le nouvel essai est un échec, soit $x_{n+1} = 0$: cela advient avec la probabilité $\omega (= 1 - \pi)$; ou elle transite vers $(0 \ 0 \ 1 \ 1)_{n+1}$ avec $x_{n+1}=1$, selon la probabilité π . Toutes les autres transitions étant impossibles, p. ex. de $(0 \ 0 \ 0 \ 0)_n$ vers $(1 \ 0 \ 1 \ 0)_{n+1}$, leurs probabilités sont nulles, et nous les avons omises pour la clarté, dans le tableau.

Structuration de la matrice absorbante T. Pour être utilisée dans le calcul de la distribution et des moments de la loi $Pa(r, k, \pi)$, la matrice élaborée ci-dessus doit être modifiée afin de la rendre « absorbante », c.-à-d. afin que son roulement, d'une transition à l'autre, respecte la règle de

couronnement, et d'arrêt, de la variable n , soit dès que $\sum_n^{(k)} x_j = r$. Pour rendre absorbante la matrice des probabilités transitionnelles $2^k \times 2^k$, il suffit de remplacer les deux sous-matrices inférieures, l'une par la matrice nulle, l'autre par la matrice identité, à l'image du modèle déjà présenté. La figure 4 présente ce résultat.

Formation du vecteur des probabilités d'entrée. Comme dernière opération préparatoire, nous devons déterminer le vecteur des probabilités d'entrée dans chaque chaîne $(x \ x \ \dots \ x)_n$, soit le vecteur $\mathbf{V}(2^k)$, dont l'arrangement doit correspondre à celui des chaînes (ou lignes) représentées dans la matrice **T** du Tableau 2.

Calcul de l'espérance et de la variance. Nous sommes prêt pour finaliser les calculs des moments et de la distribution de probabilité de la loi $Pa(4, 3, \pi)$. Fixons à présent la valeur de π , soit (par exemple) $\pi = 0,6$, avec $\omega = 1 - \pi = 0,4$.

Les moments initiaux M_i (29) sont obtenus pour $n = 3$ et 4, soit :

$$M_0 = \pi^3 + 3\omega \cdot \pi^3 = 0,4752$$

$$M_1 = \pi^3 \times 3 + 3\omega \cdot \pi^3 \times 4 = 1,6848$$

$$M_2 = \pi^3 \times 3^2 + 3\omega \cdot \pi^3 \times 4^2 = 6,0912$$

Pour trouver l'espérance (30), il nous faut d'abord, à partir de la sous-matrice **P**, former la matrice $(\mathbf{I} - \mathbf{P})$, puis inverser celle-ci pour obtenir la matrice $\mathbf{A} = (\mathbf{I} - \mathbf{P})^{-1}$: l'utilisation d'au moins 5 décimales est recommandée : les dites matrices sont présentées ci-après dans la Figure 5.

Le vecteur des probabilités d'entrée **V**, après transposition, devient :

$$\mathbf{V}^T = \{ 0,0256 \ 0,0384 \ 0,0384 \ 0,0576 \ 0,0384 \\ 0,0576 \ 0,0576 \ 0,0384 \ 0,0576 \ 0,0576 \\ 0,0576 \ 0,0864 \ 0,0864 \ 0,0864 \ 0,0864 \\ 0,1296 \}$$

Les formules (32) peuvent alors être appliquées : l'expression « $\mathbf{A} \cdot \mathbf{1}$ » dénote un vecteur-colonne obtenu par la multiplication de la matrice $\mathbf{A}_{11 \times 11}$ ci-dessus par le vecteur-colonne $\mathbf{1}_{11}$ (c.-à-d. la somme de chaque ligne de **A**). Les calculs produisent :

$$\alpha = 2,487 \ ; \ \beta = 20,440 \ .$$

Ainsi,

$$\begin{aligned} \mu = E(n) &= M_1 + k \cdot (1 - M_0) + \alpha \\ &= 1,6848 + 4 \cdot (1 - 0,4752) + 2,487 \\ &= 6,271 \end{aligned}$$

et

$$\begin{aligned} \sigma^2 = \text{Var}(n) &= M_2 + k^2 \cdot (1 - M_0) + 2k \cdot \alpha + \beta - [E(n)]^2 \\ &= 6,0912 + 42(1 - 0,4752) + 2 \times 4 \times 2,487 + 20,440 - 6,271^2 \\ &= 15,499 \ . \end{aligned}$$

Un calcul plus précis donne $\mu = 6,270997$ et $\sigma^2 = 15,498346$.

matrice **P** :

0,4	0,6									
		0,4	0,6							
				0,4	0,6					
						0,4				
							0,4	0,6		
									0,4	
										0,4
0,4	0,6									
		0,4	0,6							
				0,4	0,6					
							0,4	0,6		

matrice (**I - P**) :

0,6	-0,6									
	1	-0,4	-0,6							
		1		-0,4	-0,6					
			1			-0,4				
				1			-0,4	-0,6		
					1				-0,4	
						1				-0,4
-0,4	-0,6						1			
		-0,4	-0,6					1		
				-0,4	-0,6				1	
						-0,4	-0,6			1

Figure 5, Matrices **P**, **I - P** et **A** (début).

Calcul des probabilités $p(n)$. Sous l'horizon $n \leq k = 4$, la loi de Pascal fournit les probabilités initiales, soit :

$$p(3) = \pi^3 = 0,216$$

$$p(4) = 3\omega \cdot \pi^3 = 0,2592.$$

Viennent ensuite les calculs dérivés de la loi cumulative tels qu'exprimés dans les fonctions (34) et (35), pour $n = k + 1 = 5, 6$, etc.

Posons d'abord $Z(4) = 0$. Afin d'obtenir $p(5)$ pour $n = 5$, il nous faut $\mathbf{T}^{5-4} = \mathbf{T}^1$, soit la puissance donnée de la matrice absorbante **T**, notée à la figure 4. Nous y considérons la sous-matrice **S**, grâce à laquelle nous avons à calculer $Z(5) =$

$\mathbf{V}^T(\mathbf{S} \cdot \mathbf{1})$, c.-à-d. le produit scalaire du vecteur des probabilités d'entrée par le vecteur de la somme des lignes de **S**. Dans ce premier cas, le calcul se ramène visiblement à :

$$\begin{aligned} Z(5) &= \omega^2 \pi^2 \cdot \pi + \omega^2 \pi^2 \times \pi + \omega^2 \pi^2 \times \pi \\ &= 3\omega^2 \pi^3 \\ &= 0,10368, \end{aligned}$$

et :

$$\begin{aligned} p(5) &= Z(5) - Z(4) \\ &= 0,10368 - 0 \\ &= 0,10368. \end{aligned}$$

Afin d'obtenir ensuite $p(6)$ pour $n = 6$, la seconde

matrice $\mathbf{A} = (\mathbf{I} - \mathbf{P})^{-1}$:

1,78	1,18	0,58	0,87	0,30	0,46	0,35	0,18	0,27	0,18	0,14
0,12	1,18	0,58	0,87	0,30	0,46	0,35	0,18	0,27	0,18	0,14
0,20	0,30	1,30	0,46	0,69	1,03	0,18	0,30	0,46	0,41	0,07
0,06	0,09	0,09	1,14	0,05	0,07	0,46	0,09	0,14	0,03	0,18
0,38	0,58	0,58	0,87	1,30	0,46	0,35	0,58	0,87	0,18	0,14
0,08	0,12	0,12	0,18	0,27	1,41	0,07	0,12	0,18	0,56	0,03
0,15	0,23	0,23	0,35	0,12	0,18	1,14	0,23	0,35	0,07	0,46
0,78	1,18	0,58	0,87	0,30	0,46	0,35	1,18	0,27	0,18	0,14
0,12	0,18	0,58	0,87	0,30	0,46	0,35	0,18	1,27	0,18	0,14
0,20	0,30	0,30	0,46	0,69	1,03	0,18	0,30	0,46	1,41	0,07
0,38	0,58	0,58	0,87	0,30	0,46	0,35	0,58	0,87	0,18	0,14

Figure 5, Matrices \mathbf{P} , $\mathbf{I} - \mathbf{P}$ et \mathbf{A} (suite).

puissance de \mathbf{T} , $\mathbf{T}^{6-4} = \mathbf{T}^2$, est requise, dont on considère ensuite la sous-matrice \mathbf{S} , un même calcul fournissant $Z(6) = 0,18662$, puis $p(6) = Z(6) - Z(5) = 0,08294$. Le reste de la distribution peut s'obtenir ainsi, avec $p(7) = 0,07050$, $p(8) = 0,06304$, $p(9) = 0,04512$, etc.

Balakrishnan et Koutras (2002, p. 301 et suiv.) proposent quelques méthodes pour approximer les probabilités $p(n)$, notamment pour n fort. La modélisation des valeurs lointaines de $p(n)$ par une loi géométrique, surtout en connaissant d'avance l'espérance $E(n)$, est aussi une option (voir Laurencelle 1987).

L'estimation du paramètre π reste problématique pour le cas général de la loi de Pascal restreinte et n'a pas vraiment été étudiée, sinon solutionnée. L'estimation par inversion de l'espérance nécessite une expression pour celle-ci. Laurencelle (1998a) fournit $E(n)$ pour les lois $Pa(3, 4, \pi)$, $Pa(3, 5, \pi)$ et $Pa(4, 5, \pi)$. Une autre approche consiste, à partir d'une ou quelques réalisations n_1, n_2, \dots , à trouver par énumération les coefficients c_i de $p(n_i)$ dans l'expression générique $\sum c_i \pi^i \omega^{n-i}$ pour $i = r, r+1, \dots$ pour chaque réalisation n_j et à repérer enfin la valeur de π qui en maximise le produit. Par exemple, pour la loi $Pa(3, 4, \pi)$ illustrée ci-dessus, supposons que nous observions les réalisations $n_j = 10, 11$ et 13 . L'énumération de la combinatoire de la loi donne, pour $n = 10$, $c_i = 3, 14, 25, 5$; pour $n = 11$, $c_i = 3, 17, 39, 21, 1$; pour $n = 13$, $c_i = 3, 23, 76, 98, 40$. Pour $n = 10$, l'expression complète est donc :

$$p(10) = 3\pi^3\omega^7 + 14\pi^4\omega^6 + 25\pi^5\omega^5 + 5\pi^6\omega^4,$$

dont la valeur maximale ($\approx 0,04712$) est atteinte pour $\pi = 0,4587$; les valeurs trouvées de π pour $n_j = 11$ et 13 sont

respectivement $0,4394$ et $0,4070$. Par estimation conjointe, nous obtenons $\hat{\pi} \approx 0,4333$. Quant à l'espérance, elle est (Laurencelle 1998a) $(1 + \pi + 2\pi^2 - 3\pi^3 + 3\pi^4 - \pi^5) / [\pi^3(3 - 4\pi + 3\pi^2 - \pi^3)]$, et son inversion pour $n_j = 10, 11$ et 13 donne respectivement $0,4674, 0,4455$ et $0,4105$, alors qu'elle fournit $0,4390$ pour la valeur moyenne, $\bar{n} = 11,333$. Une dernière solution consisterait à utiliser la chaîne de Markov itérativement, soit pour l'estimation inverse à la partir de l'espérance, soit pour intercepter les probabilités qui nous concernent et les multiplier, une solution qui semble plutôt lourde avec les moyens dont nous disposons.

Pour le lecteur curieux, Laurencelle (1998b) s'est intéressé à une extension de la loi de Pascal restreinte dans le domaine du continu, en traçant le territoire des lois Gamma restreintes et en y apportant quelques solutions et plusieurs pistes d'études.

Nous fournissons, en annexe au présent article, le texte d'un programme Delphi (ou Pascal) qui effectue les calculs de l'espérance, la variance et la distribution de probabilités décrits ci-dessus. Le programme utilise une procédure d'inversion de matrice carrée, « *matinv* », inspirée d'une procédure en Fortran dans Cooley et Lohnes (1971), laquelle a le mérite discutable d'inverser la matrice fournie « sur place », sans la détruire ni recourir à une matrice auxiliaire. Nous ne garantissons pas l'entière correction de ce programme.

Références

Balakrishnan, N., Koutras, M. V. (2002). *Runs and scans with applications*. New York : Wiley.

- Bradley, J. V. (1968). *Distribution-free statistical tests*. Englewood Cliffs (NJ) : Prentice-Hall.
- Cooley, W. W., Lohnes, P. R. (1971). *Multivariate data analysis*. New York : Wiley.
- Feller, W. (1968). *An introduction to probability theory and its applications* (tome 1) (3^e édition). New York : Wiley.
- Grant, D. A. (1946). New statistical criteria for learning and problem solution in experiments involving repeated trials. *Psychological Bulletin*, 43, 272-282.
- Johnson, N. L., Kotz, S., Kemp, A. W. (1992). *Univariate discrete distributions* (2^e édition). New York : Wiley.
- Kemeny, J. G., Snell, J. L. (1976). *Finite Markov chains*. New York : Springer-Verlag.
- Koutras, M. V. (1996). On a waiting time distribution in a sequence of Bernoulli trials. *Annals of the Institute of Statistical Mathematics*, 48, 789-806.
- Koutras, M. V., Alexandrou, V. A. (1995). Runs, scans and urn model distributions : A unified Markov chain approach. *Annals of the Institute of Statistical Mathematics*, 47, 743-766.
- Laurencelle, L. (1983). La loi de Pascal restreinte. *Lettres Statistiques*, 7, 1.1-1.22.
- Laurencelle, L. (1987). La loi des succès consécutifs dans un processus de Bernoulli. *Lettres Statistiques*, 8, 25-47.
- Laurencelle, L. (1998a). La variable $n(k, 2)$ de la loi de Pascal restreinte, avec compléments. *Lettres Statistiques*, 10, 51-57.
- Laurencelle, L. (1998b). Les lois Gamma restreintes. *Lettres Statistiques*, 10, 67-84.
- Laurencelle, L. (2005). Le calcul des moments et de la distribution de probabilité de la loi de Pascal restreinte par la chaîne de Markov. *Lettres Statistiques*, 12, 65-79.
- Moivre, A. (1738). *The doctrine of chances* (2^e édition). Londres.
- Mood, A. M. (1940). The distribution theory of runs. *Annals of Mathematical Statistics*, 11, 367-392.
- Mosteller, F. (1941). Note on an application of runs to quality control charts. *Annals of Mathematical Statistics*, 12, 228-233.
- Ross, S. M. (1980). *Introduction to probability models*. New York : Academic Press.
- Uspensky, J. V. (1937). *Introduction to mathematical probability*.

Une annexe suit

Annexe

Un programme de langage Delphi (ou Pascal) pour le calcul markovien de la loi de Pascal restreinte

```

program pascal_restreinte;
  {$APPTYPE CONSOLE }
  { L. Laurencelle, septembre 2011
    d'après Lettres Statistiques, 2005 }
uses SysUtils, Math;

const maxo=1024;

type matrice=array[1..maxo,1..maxo] of double;
  vecteuri=array[1..maxo] of integer;
  vecteurr=array[1..maxo] of double;

var k,r,ordre:integer; pi,w : double;
  n,nmax:integer;
  i,j,L,masque,ordrep1,c,b,h,t,x,y0,y1:integer;
  mat1,mat2,mat3:matrice;
  liste:vecteuri;pliste,rmat3:vecteurr;
  Pind,Ptot,Pmark,np1,vp1,np2,vp2,esp,vari,det,v,cnp2,cvp2:double;

procedure matinv(var A:matrice; var M:integer; var DET:double);
var j,k,L:integer;pvt,tt:double;
begin
  det:=1;
  for j:=1 to M do
  begin pvt:=A[j,j];
    det:=det*pvt;
    A[j,j]:=1;
    for k:=1 to M do A[j,k]:=A[j,k]/pvt;
    for k:=1 to M do
    if k<>j then begin tt:=A[k,j]; A[k,j]:=0;
      for L:=1 to M do A[k,L]:=A[k,L]-A[j,L]*tt
    end
  end
end;

begin
repeat
{ cueillir les parametres et tout initialiser : k < 12 }
  write(' r k P : ');readln(r,k,Pi);
  w:=1-Pi;
  ordre:=2;for i:=2 to k do ordre:=ordre*2;
  masque:=ordre-1;
  ordrep1:=ordre+1;
{ former la matrice absorbante T, `d'ordre 2^k x 2^k }
  b:=0;h:=0;
  for x:=0 to ordre-1 do
  begin c:=0; t:=x;
    while t>0 do begin if (t mod 2)=1 then c:=c+1;t:=t div 2 end;
    if c<r then
    begin b:=b+1; liste[b]:=x; pliste[b]:=power(pi,c)*power(w,k-c) end
    else
    begin h:=h+1; liste[ordrep1-h]:=x
    end
  end;
  for i:=1 to ordre do for j:=1 to ordre do begin mat1[i,j]:=0; mat2[i,j]:=0 end;
  for i:=ordrep1-h to ordre do begin mat1[i,i]:=1; mat2[i,i]:=1 end;
  for i:=1 to b do
  begin
    y0:=2*liste[i] and masque;
    j:=1;

```

```

while liste[j]<>y0 do j:=j+1;
mat1[i,j]:=w;
y1:=y0+1;j:=j+1;
while liste[j]<>y1 do j:=j+1;
mat1[i,j]:=pi
end;
for i:=1 to b do mat2[i,i]:=1;
{ calculer la premiere phase de la distribution,
pour n <= k, ainsi que la premiere phase de l'esperance E(n)
et de la variance Var(n) }
Pind:=power(pi,r);
np1:=Pind*r; vp1:=Pind*sqr(r);
Ptot:=Pind;
writeln(r:3,Pind:9:6,Ptot:9:6);
if k>r then
  for n:=r+1 to k do
  begin
  Pind:=Pind*w*(n-1)/(n-r);
  np1:=np1+Pind*n; vp1:=vp1+Pind*sqr(n);
  Ptot:=Ptot+Pind;
  writeln(n:3,Pind:9:6,Ptot:9:6)
  end;
{ inverser (I - P), P etant la sous-matrice de pietinements c x c, ou c < r,
pour en obtenir la 2e phase de l'esperance et la variance }
for i:=1 to b do for j:=1 to b do
  if i=j then mat3[i,j]:=1-mat1[i,j] else mat3[i,j]:=-mat1[i,j];
matinv(mat3,b,det);
np2:=0;
for i:=1 to b do
begin v:=0;
  for j:=1 to b do v:=v+mat3[i,j];
  rmat3[i]:=v;
  np2:=np2+v*pliste[i]
end;
vp2:=0;
for i:=1 to b do
  begin v:=0;
  for j:=1 to b do
    if i=j then v:=v+(2*mat3[i,j]-1)*rmat3[j]
    else v:=v+(2*mat3[i,j] )*rmat3[j];
  vp2:=vp2+v*pliste[i]
  end;
{ calculer et ajouter le decalage (par rapport a l'entree dans la roue
markovienne) a partir de n = k+1 }
cnp2:=k*(1-Ptot);
cvp2:=sqr(k)*(1-Ptot);
Esp:=np1+(np2+cnp2);
Vari:=vp1+cvp2+2*k*np2+vp2 - sqr(Esp);
{ calculer les probabilites p(n) par les puissances successives de T, ce
jusqu'a un critere determine (selon P ou n) : le critere ici applique
est Ptot >0.999. Le calcul est dedouble afin de faire alterner les
matrices multiplicande et produit }
n:=k;
Pmark:=0;
while Ptot<0.999 do {CONDITION D'ARRET}
{ premier cycle de calcul }
begin
  n:=n+1;
  for i:=1 to ordre do for j:=1 to ordre do
  begin v:=0;
    for L:=1 to ordre do v:=v+mat1[i,L]*mat2[L,j];
    mat3[i,j]:=v
  end;
  Pind:=0;
  for i:=1 to b do
  begin v:=0;

```



```

    for j:=ordrepl-h to ordre do v:=v+mat3[i,j];
    Pind:=Pind+Pliste[i]*v
end;
Pind:=Pind-PMark;
Pmark:=Pmark+Pind;
Ptot:=Ptot+Pind;
writeln(n:3,Pind:9:6,Ptot:9:6);
{ second cycle de calcul }
n:=n+1;
for i:=1 to ordre do for j:=1 to ordre do
begin v:=0;
  for L:=1 to ordre do v:=v+mat1[i,L]*mat3[L,j];
  mat2[i,j]:=v
end;
Pind:=0;
for i:=1 to b do
begin v:=0;
  for j:=ordrepl-h to ordre do v:=v+mat2[i,j];
  Pind:=Pind+Pliste[i]*v
end;
Pind:=Pind-PMark;
Pmark:=Pmark+Pind;
Ptot:=Ptot+Pind;
writeln(n:3,Pind:9:6,Ptot:9:6);
end;
{ terminer }
writeln('Esperance =',Esp:10:6,' Variance =',Vari:12:6)
until false
end.

```