JOURNAL

DE

MATHÉMATIQUES

PURES ET APPLIQUÉES

FONDÉ EN 1836 ET PUBLIÉ JUSQU'EN 1874

PAR JOSEPH LIOUVILLE

E. FELDHEIM

Nouvelle démonstration et généralisation d'un théorème du calcul des probabilités dû à Simmons

Journal de mathématiques pures et appliquées 9e série, tome 20 (1941), p. 1-16.

<http://www.numdam.org/item?id=JMPA_1941_9_20__1_0>

gallica NUMDAM

Article numérisé dans le cadre du programme Gallica de la Bibliothèque nationale de France
http://gallica.bnf.fr/

et catalogué par Mathdoc dans le cadre du pôle associé BnF/Mathdoc
http://www.numdam.org/journals/JMPA

JOURNAL
DE
MATHÉMATIQUES
PURES ET APPLIQUÉES.

Nouvelle démonstration et généralisation d'un théorème du calcul des probabilités dû à Simmons;

Par E. FELDHEIM.

1. Introduction. — Considérons une urne contenant a boules noires et b boules blanches ($a > b$), et faisons $n = N(a+b)$ extractions, en remettant toujours dans l'urne la boule tirée. Le nombre des boules blanches extraites au cours des n tirages sera désigné par S. On a le théorème suivant :

Il est plus probable que S est inférieur plutôt que supérieur au nombre Nb, c'est-à-dire que, si l'on désigne les probabilités des inégalités $S < Nb$ et $S > Nb$ par P et Q respectivement, on aura

$$P > Q, \qquad \text{pourvu que} \quad a > b.$$

Ensuite, si N est assez grand, la différence $P - Q$ est approximativement égale à l'expression

$$\Delta = \frac{a-b}{3\sqrt{2\pi nab}}.$$

Le théorème précédent, avec une interprétation différente de celle que nous en avons donnée, a été démontré par Simmons dans un Mémoire de 43 pages ([1]), où il l'a complété de remarques expérimentales sur les jeux de hasard, etc. Ses résultats, malgré leur importance et simplicité, sont assez peu connus. Nous avons donc cru utile de donner une nouvelle démonstration pour ce théorème, en l'étendant au cas plus général des épreuves répétées.

Notre travail consiste en trois Chapitres. Dans le premier, après avoir rappelé quelques propriétés des probabilités des épreuves répétées, importantes pour la suite du travail, nous donnons la démonstration de l'inégalité $P > Q$, et nous traitons des questions attachées à ce problème.

Le second Chapitre contient la démonstration de la seconde partie du théorème, c'est-à-dire la détermination de la valeur asymptotique de la différence $P - Q$. Ce calcul est effectué par une simplification notable de la méthode de Simmons.

Revenons maintenant à l'énoncé simple ci-dessus du théorème. Désignons par p la probabilité du tirage d'une boule blanche

$$p = \frac{b}{a+b}.$$

Nous avons alors

$$Nb = np,$$

Nb est donc égal à la moyenne arithmétique (ou valeur moyenne) du nombre des boules blanches. Le *nombre np est ainsi un nombre entier*. La démonstration du théorème, donnée dans les deux Chapitres premiers, n'est donc valable que pour ce cas particulier.

Dans le troisième Chapitre, nous démontrons le théorème en toute sa généralité, rapporté au cas des épreuves répétées, et faisant abstraction de la condition restrictive du Chapitre précédent. L'idée principale

---

([1]) T. C. Simmons, *A New Theorem of Probability* (*Proceed. of the London Math. Soc.*, vol. 26, 1ère série, 1894-1895, p. 290-334). Parmi les travaux s'occupant encore du théorème de Simmons, *voir :* Ragnar Frisch, *Solution d'un problème du calcul des probabilités. Premier problème de Simmons* (*Skandinavisk Aktuarietidskrift*, 1924, p. 153-174, et *C. R. Acad. Sc. Paris*, nov. 1924).

de cette démonstration est analogue à celle de la loi de Laplace au moyen des probabilités des épreuves répétées.

2. Démonstration de l'inégalité $P > Q$. — Soit p la probabilité constante d'un événement E, et q celle de l'événement contraire E'. La probabilité pour que, au cours de n expériences indépendantes, la répétition de E soit r, celle de E' soit $n-r$, est égale, d'après les résultats classiques, à

$$T_r = \binom{n}{r} p^r q^{n-r} \qquad (p+q=1). \tag{1}$$

Mentionnons quelques propriétés, bien connues d'ailleurs, de ces probabilités, dont nous aurons besoin dans la suite.

a. La probabilité T_r est le terme d'ordre $r+1$ du développement en série du binome $(p+q)^n$.

b. L'indice r du plus grand terme du développement satisfait aux inégalités $np-q < r \leqq np+p$. Nous pouvons dire que T_r est maximum lorsque l'indice r est l'entier le plus voisin du nombre np. Si l'on introduit la *fréquence* $\frac{r}{n}$ de l'événement en question, nous pouvons aussi dire que la fréquence la plus probable est celle qui est la plus voisine de la probabilité. Si np est entier, ce que nous supposons dans la première partie du travail, la fréquence la plus probable est égale à la probabilité, et T_r est maximum pour $r = np = m$.

Si nous considérons la probabilité T_r comme une fonction de p, et si nous cherchons la valeur de p pour laquelle cette fonction est maximum, c'est-à-dire $\frac{dT_r}{dp} = 0$, nous trouvons encore la relation $r = np$.

Cherchons maintenant la valeur moyenne des valeurs de r

$$m = \sum_{r=0}^{n} r T_r = np. \tag{2}$$

Si np est un nombre entier, la valeur moyenne des r coïncide avec l'indice pour lequel T_r est maximum.

c. La probabilité de la combinaison la plus probable est donc approximativement (et, si np est entier, alors exactement) égale à

$$T_m = \binom{n}{m} p^{np} q^{nq} \qquad (m = np).$$

L'application de la formule de Stirling donne, en cas de n assez grand, la valeur approchée suivante pour T_m

$$T_m \sim \frac{1}{\sqrt{2\pi npq}}. \tag{3}$$

Si $n \to \infty$, la probabilité T_m tend vers zéro.

La première partie du théorème affirme que, dans le cas où $p < q$, la probabilité des valeurs de r plus petites que la moyenne (où l'indice du maximum) est plus grande que la probabilité des valeurs de r plus grandes que la moyenne (ou l'indice du maximum), ou autrement, qu'une déviation négative $r - np$ est plus probable qu'une déviation positive. Si la probabilité de la vérification de l'événement E est Prob $\{E\}$, on a

$$P = \text{Prob}\{r < m\}, \qquad Q = \text{Prob}\{r > m\},$$

notre théorème s'énonce alors de la façon suivante :

L'inégalité $p < q$ entraîne l'inégalité $P > Q$.

Avec les notations précédentes,

$$P = \sum_{r=0}^{m-1} T_r, \qquad Q = \sum_{r=m+1}^{n} T_r. \tag{4}$$

Nous avons besoin pour la démonstration de l'inégalité $P > Q$ d'un résultat auxiliaire. En tenant compte de ce que $\sum_{r=0}^{n} T_r = 1$, on déduit de (2) la relation

$$\sum_{r=0}^{n} (m - r) T_r = 0,$$

c'est-à-dire

$$\sum_{r=0}^{m-1} (m - r) T_r = \sum_{r=m+1}^{n} (r - m) T_r. \tag{5}$$

Posons maintenant

$$T_{m-r} = T_m u_r, \qquad \text{où} \quad r = 1, 2, 3, \ldots, m$$

et

$$T_{m+r} = T_m v_r, \qquad \text{où} \quad r = 1, 2, 3, \ldots, n-m.$$

Les relations (4) prennent alors la forme

$$(6) \qquad P = T_m \sum_{r=1}^{m} u_r, \qquad Q = T_m \sum_{r=1}^{n-m} v_r,$$

et, d'après (5), on a

$$(7) \qquad \sum_{r=1}^{m} r u_r = \sum_{r=1}^{n-m} r v_r.$$

Désignons par k le premier indice pour lequel

$$u_k < v_k.$$

Ce nombre k est nécessairement plus grand que 1, parce que

$$u_1 = \frac{1}{1 + \frac{1}{nq}}, \qquad v_1 = \frac{1}{1 + \frac{1}{np}},$$

et l'on voit immédiatement que $u_1 > v_1$, si $p < q$.

On peut alors écrire les inégalités suivantes

$$(8) \qquad \sum_{r=k}^{m} (r - k + 1) u_r < \sum_{r=k}^{n-m} (r - k + 1) v_r$$

et

$$(9) \qquad \sum_{r=1}^{k-1} (k - r - 1) u_r > \sum_{r=1}^{k-1} (k - r - 1) v_r.$$

Formons la différence de (7) et de (8), et ajoutons-y l'inégalité (9)

$$(k-1) \sum_{r=1}^{m} u_r > (k-1) \sum_{r=1}^{n-m} v_r.$$

Nous avons vu que $k > 1$, et ainsi

$$\sum_{r=1}^{m} u_r > \sum_{r=1}^{n-m} v_r,$$

ce qui est identique à l'inégalité que nous voulions démontrer.

Faisons encore une remarque sur la valeur de k. La démonstration précédente est valable pour $k \leqq m$. Si l'on a $u_r > v_r$ pour toute valeur possible de r : $r \leqq m$, on fait $k = m + 1$, le premier membre de l'inégalité (8) est remplacé par o, et le reste de la démonstration ne change pas.

Revenons maintenant à la relation (5), au sujet de laquelle nous pouvons faire quelques remarques intéressantes.

Désignons par m_1 et m_2 les valeurs moyennes des valeurs de r plus petites, ou respectivement plus grandes que la moyenne de r, c'est-à-dire

$$m_1 = \frac{1}{P} \sum_{r=1}^{m} r T_r = \frac{m}{P} \sum_{r=0}^{m-2} \binom{n-1}{r} p^r q^{n-1-r} = \frac{m P'}{P};$$

$$m_2 = \frac{1}{Q} \sum_{r=m+1}^{n} r T_r = \frac{m}{Q} \sum_{r=m}^{n-1} \binom{n-1}{r} p^r q^{n-1-r} = \frac{m Q'}{Q},$$

d'où

$$m(P - P') = P(m - m_1) = \sum_{r=0}^{m-1} (m - r) T_r$$

et

$$m(Q' - Q) = Q(m_2 - m) = \sum_{r=m+1}^{n} (r - m) T_r.$$

On a donc, d'après (5),

$$P - P' = Q' - Q. \tag{10}$$

Nous obtenons encore facilement un résultat de A. Guldberg [2]. Il résulte de (10), et des formules précédentes, que

$$P(m - m_1) = Q(m_2 - m),$$

---

(2) *Nouv. Ann. de Math.*, 1922-1923, p. 251.

d'où

$$P \gtreqless Q, \qquad \text{si l'on a respectivement} \quad m \gtreqless \frac{m_1 + m_2}{2}.$$

On peut aussi définir très simplement la valeur commune des différences (10). La relation connue

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1} \tag{11}$$

permet d'écrire que

$$P = q \sum_{r=0}^{m-1} \binom{n-1}{r} p^r q^{n-1-r} + p \sum_{r=0}^{m-2} \binom{n-1}{r} p^r q^{n-1-r} = q \binom{n-1}{m-1} p^{m-1} q^{n-m}$$
$$+ (p+q) \sum_{r=0}^{m-2} \binom{n-1}{r} p^r q^{n-1-r} = q T_m + P',$$

donc la valeur commune des différences (10) est qT_m, et celle des expressions (7) est mq.

Remarquons encore que la valeur commune de $P - P'$ et $Q' - Q$ s'obtient aussi comme un cas particulier d'une formule de Bertrand ([3]). D'après cette formule

$$\sum_{r=0}^{\rho} (np - r) T_r + \sum_{r=\rho+1}^{n} (r - np) T_r = 2p(n-\rho) T_\rho, \tag{12}$$

où ρ désigne un nombre entier. Si $np = m$ est entier, et $\rho = m$, le premier membre de (12) est identique, d'après ce qui précède, à

$$m(P - P') + m(Q' - Q) = 2m(P - P'),$$

de sorte que

$$m(P - P') = p(n - m) T_m = mq T_m,$$

c'est-à-dire le résultat établi tout à l'heure.

Dans le cas général, la relation (12) s'écrit de la façon suivante

$$\left(P - P' + \frac{np - \rho}{np} T_\rho\right) + (Q' - Q) = 2 \frac{n - \rho}{n} T_\rho;$$

([3]) Pour une démonstration récente, *voir* M. Fréchet, *Sull'espressione esatta di uno scarto medio* (*Giornale dell'Istituto Italiano degli Attuari*, vol. **6**, 1935).

d'autre part, il est évident que

$$P+Q=1-T_\rho \qquad \text{et} \qquad P'+Q'=1-\binom{n-1}{\rho-1}p^{\rho-1}q^{n-\rho}=1-\frac{\rho}{np}T_\rho,$$

d'où

$$P-P'+\frac{np-\rho}{np}T_\rho=Q'-Q,$$

donc

$$P-P'=\frac{q}{p}\frac{\rho}{n}T_\rho, \qquad Q'-Q=\left(1-\frac{\rho}{n}\right)T_\rho$$

pour toute valeur de l'entier ρ.

3. Détermination de la valeur asymptotique de la différence $P-Q$. — Nous pouvons compléter l'inégalité $P>Q$ démontrée dans **2**, par la valeur asymptotique de $P-Q$ donnée par Simmons, qui prend, avec nos notations, la forme suivante

$$\Delta=\frac{q-p}{3\sqrt{2\pi npq}}.$$

La démonstration sera effectuée à l'aide de transformations identiques du développement de la série de binome. On se bornera le plus souvent à des développements formels, et l'on fait aussi abstraction de la signification en calcul des probabilités des expressions qui figureront dans la suite. Nous ne devons en tenir évidemment compte que pour les calculs, les résultats, à part de la restriction déjà mentionnée, étant tout à fait généraux.

On peut démontrer, au moyen de la formule (11), les relations suivantes

$$(13) \qquad \sum_{r=0}^{m}\binom{n+1}{r}p^r q^{n+1-r}-\sum_{r=0}^{m-1}\binom{n}{r}p^r q^{n-r}=qT_m,$$

$$(14) \qquad \sum_{r=0}^{m}\binom{n+k+1}{r}p^r q^{n+k+1-r}-\sum_{r=0}^{m}\binom{n+k}{r}p^r q^{n+k-r}=-p\binom{n+k}{m}p^m q^{n+k-m}$$

$$=-\left(\frac{nq}{n-m}\right)^k\frac{\left(1+\frac{1}{n}\right)\left(1+\frac{2}{n}\right)\cdots\left(1+\frac{k}{n}\right)}{\left(1+\frac{1}{n-m}\right)\left(1+\frac{2}{n-m}\right)\cdots\left(1+\frac{k}{n-m}\right)}pT_m.$$

Cette relation est valable pour toute valeur entière de m et n; pour $k=0$ on obtient, par une petite transformation, la formule (13). Soit successivement $k=0, 1, 2, \ldots, s$, et prenons la somme des relations (14) correspondantes

$$\sum_{r=0}^{m}\binom{n+s+1}{r}p^r q^{n+s+1-r}-\sum_{r=0}^{m-1}\binom{n}{r}p^r q^{n-r}$$
$$=q\mathrm{T}_m-p\mathrm{T}_m\sum_{r=1}^{s}\frac{\left(1+\frac{1}{n}\right)\left(1+\frac{2}{n}\right)\cdots\left(1+\frac{r}{n}\right)}{\left(1+\frac{1}{nq}\right)\left(1+\frac{2}{nq}\right)\cdots\left(1+\frac{r}{nq}\right)},$$

où $m=np$. Cherchons le développement du second membre en série de puissances croissantes de $\frac{1}{n}$. Le résultat du calcul est le suivant

$$(15)\quad \sum_{r=0}^{m}\binom{n+s+1}{r}p^r q^{n+s+1-r}-\sum_{r=0}^{m-1}\binom{n}{r}p^r q^{n-r}$$
$$=(q-sp)\mathrm{T}_m$$
$$+\frac{p^2 s(s+1)(s+2)}{6nq}\left[1-\frac{3ps^2+(p+10)s+10-4p}{20nq}+\ldots\right]\mathrm{T}_m.$$

Cette relation reste valable si l'on donne à s toute valeur entière ou fractionnaire.

Il résulte de $p+q=1$ que l'on a

$$n=np\left(1+\frac{q}{p}\right),$$

c'est-à-dire, si $s=\frac{q}{p}$, alors

$$n=m(s+1),\qquad n+s+1=(m+1)(s+1).$$

Si l'on remplace dans (15) p et q par $\frac{1}{s+1}$ et $\frac{s}{s+1}$, et si l'on considère s comme un entier, on obtient le cas particulier traité par Simmons. En substituant, après avoir effectué les calculs, $\frac{a}{b}$ au lieu de s, on aura finalement le résultat énoncé dans l'Introduction. Nous pouvons le faire en une seule fois, si nous remplaçons d'après ce qui précède, dans (15), s par le rapport $\frac{q}{p}$, n et $n+s+1$ par les valeurs calculées. Si nous désignons alors, pour simplifier l'écriture, $\sum_{r=0}^{m-1}\binom{n}{r}p^r q^{n-r}$ par P_m, le premier membre de (15) sera P_{m+1}, et (15) prend ainsi la forme suivante

$$(16)\qquad P_{m+1}-P_m=\frac{p+1}{6m}\left(1-\frac{6+9q-2q^2}{20mq}+\ldots\right)T_m \qquad (m=np).$$

Il résulte de la relation $P+Q+T_m=1$ que

$$D=P-Q=P_m-Q_m=2P_m+T_m-1=D_m,$$

d'où

$$(17)\qquad D_{m+1}-D_m=2(P_{m+1}-P_m)+(T_{m+1}-T_m).$$

Mais

$$\frac{T_{m+1}}{T_m}=\frac{\left(1+\frac{1}{n}\right)\left(1+\frac{2}{n}\right)\cdots\left(1+\frac{k}{n}\right)}{\left(1+\frac{1}{nq}\right)\left(1+\frac{2}{nq}\right)\cdots\left(1+\frac{k}{nq}\right)}, \qquad \text{où}\quad k=\frac{q}{p}.$$

Un calcul analogue au précédent donne que

$$(18)\qquad T_{m+1}=\left(1-\frac{1}{2m}+\frac{2+7q+2q^2}{24m^2q}+\ldots\right)T_m,$$

de sorte que, d'après la formule (17),

$$(19)\qquad D_{m+1}-D_m=-\frac{q-p}{6m}\left(1+\frac{2q^2-17q-14}{20mq}+\ldots\right)T_m.$$

Il est évident que, pour $n\to\infty$ (ou, ce qui revient au même, $m\to\infty$), $D\to 0$; on peut donc écrire que

$$(20)\qquad D_m=\left(X+\frac{Y}{m}+\frac{Z}{m^2}+\ldots\right)T_m,$$

où les coefficients du développement sont inconnus. Étant donné que nous cherchons seulement une valeur asymptotique pour D, il nous suffira de connaître les premiers coefficients du développement précédent. Or

$$\frac{Y}{m+1}=\frac{Y}{m}-\frac{Y}{m^2}+\ldots \quad \text{et} \quad \frac{Z}{(m+1)^2}=\frac{Z}{m^2}-\frac{2Z}{m^3}+\ldots,$$

donc,

$$\begin{aligned}D_{m+1}&=\left(X+\frac{Y}{m}+\frac{Z-Y}{m^2}+\ldots\right)T_{m+1}\\&=\left(X+\frac{2Y-X}{2m}+\frac{12q(2Z-3Y)+X(2+7q+2q^2)}{24m^2q}+\ldots\right)T_m,\end{aligned}$$

d'où

$$D_{m+1}-D_m=\left(-\frac{X}{2m}+\frac{X(2+7q+2q^2)-36qY}{24m^2q}+\ldots\right)T_m. \tag{21}$$

En comparant les coefficients des puissances égales de $\frac{1}{m}$ dans les développements (19) et (21), on trouve pour X et Y les valeurs

$$X=\frac{q-p}{3}, \qquad Y=-\frac{q-p}{3}\,\frac{4(q+1)(2-q)}{45q}.$$

On déduit maintenant de la relation (20) que

$$P-Q=\frac{q-p}{3}\left(1-\frac{4(q+1)(2-q)}{45npq}+\ldots\right)T_m. \tag{22}$$

Si $n\to\infty$, en tenant compte de la relation (3), la différence D aura la valeur asymptotique suivante

$$\Delta=\frac{q-p}{3\sqrt{2\pi npq}}. \tag{23}$$

Remarque. — On aurait pu trouver, d'une façon analogue à celle qui a conduit à la formule (20), le développement de la probabilité P suivant les puissances de $\frac{1}{m}$, en partant de la relation (16), et en déduire ensuite le développement de D. Inversement, on peut écrire, au moyen de (22), la probabilité P sous la forme suivante

$$P=\frac{1}{2}-\frac{p+1}{3}\left(1-\frac{2(q+1)(2q-1)}{45npq}+\ldots\right)T_m. \tag{24}$$

Si n devient infiniment grand, P (et Q aussi) tend vers $\frac{1}{2}$, ce qui était d'ailleurs à prévoir.

4. Étude du cas général. — Dans cette partie de notre Mémoire, nous ferons abstraction de la condition restrictive que le produit np est un nombre entier, et nous désignerons par m la partie entière de np

$$m = np - s \qquad \text{où} \quad 0 \leqq s \leqq 1.$$

La méthode que nous allons employer est, comme nous l'avons déjà indiqué dans l'Introduction, analogue à celle qui permet de déduire le théorème de Bernoulli des probabilités des épreuves répétées.

Nous aurons besoin tout d'abord d'une représentation intégrale des probabilités P et Q que nous allons chercher dans ce qui suit. D'après les formules (4),

$$(25) \qquad \begin{cases} P = \displaystyle\sum_{r=0}^{m-1} \binom{n}{r} p^r (1-p)^{n-r} = \sum_{r=0}^{m-1} \binom{n}{r} q^{n-r} (1-q)^r, \\ Q = \displaystyle\sum_{r=m+1}^{n} \binom{n}{r} p^r (1-p)^{n-r} = \sum_{r=0}^{n-m-1} \binom{n}{r} p^{n-r} (1-p)^r. \end{cases}$$

Si l'on fait

$$(26) \qquad F_k(x) = \sum_{r=0}^{k} \binom{n}{r} x^{n-r} (1-x)^r,$$

on aura donc

$$P = F_{m-1}(q), \qquad Q = F_{n-m-1}(p).$$

Dérivons les deux membres de l'équation (26) par rapport à x,

$$\begin{aligned} F'_k(x) &= \sum_{r=0}^{k} \binom{n}{r} \Big[(n-r)\, x^{n-r-1} (1-x)^r - r x^{n-r} (1-x)^{r-1} \Big] \\ &= \sum_{r=0}^{k} \frac{n!}{r!\,(n-r-1)!} x^{n-r-1} (1-x)^r - \sum_{r=1}^{k} \frac{n!}{(r-1)!\,(n-r)!} x^{n-r} (1-x)^{r-1} \\ &= \sum_{r=0}^{k} \frac{n!}{r!\,(n-r-1)!} x^{n-r-1} (1-x)^r - \sum_{r=0}^{k-1} \frac{n!}{r!\,(n-r-1)!} x^{n-r-1} (1-x)^r \\ &= \frac{n!}{k!\,(n-k-1)!} x^{n-k-1} (1-x)^k. \end{aligned}$$

On a donc, en intégrant [4]

$$F_k(x) = (n-k)\binom{n}{k}\int_0^x t^{n-k-1}(1-t)^k dt, \tag{27}$$

et ainsi

$$\begin{cases} P = m\binom{n}{m}\int_0^q t^{n-m}(1-t)^{m-1} dt, \\ Q = (n-m)\binom{n}{m}\int_0^p t^m(1-t)^{n-m-1} dt. \end{cases} \tag{28}$$

Si n est assez grand, la valeur asymptotique de l'intégrale

$$B = \frac{n!}{h!\,l!}\int_0^x t^l(1-t)^h dt, \qquad h+l = n-1, \tag{29}$$

s'obtient par l'introduction de la nouvelle variable

$$t = \frac{l}{n-1} - v\,\frac{1}{n-1}\sqrt{\frac{hl}{n-1}}$$

et en développant la fonction à intégrer (ou plutôt son logarithme) en une série de Taylor. Le résultat est le suivant [5]

$$B = \frac{1}{\sqrt{2\pi}}\int_T^\infty e^{-\frac{t^2}{2}} dt = \frac{1}{2} - \frac{1}{2}\Phi(T), \tag{30}$$

---

(4) Observons que, d'après une communication aimable de M. le prof. Ch. Jordan, des propriétés connues des fonctions Bêta incomplètes permettent de retrouver la forme (26) des probabilités $F_k(x)$, en transformant la relation (27) par une série d'intégrations par parties. Si l'on fait

$$I_x(n-k, k+1) = \frac{1}{B(n-k, k+1)}\int_0^x t^{n-k-1}(1-t)^k dt,$$

et si k et n sont des entiers, on démontre que $F_k(x)$ est égale à la valeur précédente. Pour le calcul numérique de ces intégrales, *voir* PEARSON, *Table of the incomplete Bêta-function* (*Cambridge University Press*, 1934).

(5) Pour les détails de ce calcul, *voir* B. L. VAN DER WAERDEN, *Empirische Bestimmung von Wahrscheinlichkeiten*, usw. (*Berichten der Akad. zu Leipzig*, vol. 87, 1935, p. 353-364).

où $\Phi(T)$ est l'intégrale de probabilité bien connue

$$(31)\qquad \Phi(T) = \frac{2}{\sqrt{2\pi}} \int_0^T e^{-\frac{t^2}{2}} dt.$$

Ici

$$(32)\qquad T = C + D + O\left(n^{-\frac{3}{2}}\right),$$

et

$$(33)\qquad \begin{cases} C = [l - x(n-1)]\sqrt{\dfrac{n-1}{hl}}, \\ D = \dfrac{1}{3}\left(\dfrac{13}{12(n-1)} - \dfrac{n-1}{12hl}\right)(C^3 + 3C) \\ \qquad - \dfrac{l-h}{3\sqrt{hl(n-1)}}(C^2+2) - \dfrac{(l-h)^2}{18hl(n-1)}C^5. \end{cases}$$

C'est cette formule que nous allons appliquer pour l'évaluation de la valeur asymptotique de la différence $P - Q$. En désignant les valeurs de T correspondant à P et Q respectivement par T_1 et T_2, nous aurons

$$P = \frac{1}{\sqrt{2\pi}} \int_{T_1}^{\infty} e^{-\frac{t^2}{2}} dt, \qquad Q = \frac{1}{\sqrt{2\pi}} \int_{T_2}^{\infty} e^{-\frac{t^2}{2}} dt,$$

et

$$(34)\qquad \Delta = P - Q = \frac{1}{\sqrt{2\pi}} \int_{T_1}^{T_2} e^{-\frac{t^2}{2}} dt = \frac{1}{2}[\Phi(T_2) - \Phi(T_1)].$$

Comme nous l'avons vu, D est de l'ordre de $n^{-\frac{1}{2}}$, la différence Δ peut donc être écrite de la façon suivante

$$(35)\qquad \Delta = \frac{1}{\sqrt{2\pi}}(T_2 - T_1) - \frac{1}{3\sqrt{2\pi}}(T_2^3 - T_1^3) + \ldots = \frac{1}{\sqrt{2\pi}}(T_2 - T_1) + O\left(n^{-\frac{3}{2}}\right).$$

Les valeurs T_1 et T_2 qui y figurent sont calculées à l'aide de (32), où l'on a, dans le cas de T_1 :

$$l_1 = nq + s, \qquad h_1 = np - s - 1, \qquad x_1 = q;$$

dans le cas de T_2 :

$$l_2 = np - s, \qquad h_2 = nq + s - 1, \qquad x_2 = p.$$

Alors

$$C_1 = (s+q)\sqrt{\frac{n-1}{(nq+s)(np-s-1)}}.$$

L'expression sous le signe de racine carrée se met sous la forme

$$\frac{1}{\sqrt{npq}}\left(1+\frac{s(q-p)+q^2}{2npq}+\ldots\right),$$

donc

$$C_1 = \frac{s+q}{\sqrt{npq}} + O\left(n^{-\frac{3}{2}}\right).$$

Dans l'expression de D_1, le premier terme est de l'ordre de $n^{-\frac{3}{2}}$, le troisième de l'ordre $n^{-\frac{7}{2}}$. Dans le second terme, l'ordre de $C_1^2 \frac{l_1-h_1}{3\sqrt{(n-1)h_1 l_1}}$ est $n^{-\frac{3}{2}}$, il suffit donc de considérer le terme $-\frac{2(l_1-h_1)}{3\sqrt{(n-1)l_1 h_1}}$,

$$D_1 = -\frac{2(q-p)}{3\sqrt{npq}}\left(1+\frac{s+q^2(3-2q)}{2npq(q-p)}+\ldots\right)+O\left(n^{-\frac{3}{2}}\right).$$

Enfin

$$(36)\qquad T_1 = \frac{3(s+q)-2(q-p)}{3\sqrt{npq}} + O\left(n^{-\frac{3}{2}}\right) = \frac{3s+p+1}{3\sqrt{npq}} + O\left(n^{-\frac{3}{2}}\right).$$

Un calcul tout à fait analogue au précédent donne que

$$(37)\qquad T_2 = \frac{-3s+q+1}{3\sqrt{npq}} + O\left(n^{-\frac{3}{2}}\right).$$

On voit sur l'expression (35) de la différence Δ, que $\Delta > 0$ (c'est-à-dire $P > Q$), si $T_2 > T_1$. Dans ce cas, la condition de cette inégalité est que

$$-3s+q+1 > 3s+p+1,$$

d'où

$$6s < q-p.$$

Nous avons, par hypothèse, $q > p$, et ainsi, d'après l'inégalité précédente,

$$0 < s < \frac{q-p}{6},$$

c'est-à-dire, le nombre entier m doit satisfaire à l'inégalité

$$np - \frac{q-p}{6} \leqq m \leqq np,$$

pour que l'inégalité $T_2 > T_1$ soit vérifiée. Remarquons toutefois que nous avons supposé le nombre n assez grand.

La valeur asymptotique de la différence Δ est

$$\Delta_s = \frac{q-p-6s}{3\sqrt{2\pi npq}} + O\left(n^{-\frac{3}{2}}\right). \tag{38}$$

Dans le cas particulier où le produit np est un nombre entier, c'est-à-dire où $s = 0$, la valeur asymptotique de la différence $P - Q$ est identique à l'expression trouvée dans le paragraphe précédent

$$\Delta_0 = \frac{q-p}{3\sqrt{2\pi npq}}.$$