

Zeitschrift: Technische Mitteilungen / Schweizerische Post-, Telefon- und Telegrafienbetriebe = Bulletin technique / Entreprise des postes, téléphones et télégraphes suisses = Bollettino tecnico / Azienda delle poste, dei telefoni e dei telegrafi svizzeri

Herausgeber: Schweizerische Post-, Telefon- und Telegrafienbetriebe

Band: 40 (1962)

Heft: 6

Artikel: Méthode de détermination de la probabilité d'exactitude d'un résultat par sondages et étude comparative avec la méthode de Copper et Pearson

Autor: Wohlers, R.

DOI: <https://doi.org/10.5169/seals-875123>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

Download PDF: 30.04.2026

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Méthode de détermination de la probabilité d'exactitude d'un résultat obtenu par sondages et étude comparative avec la méthode de Clopper et Pearson

Le problème à résoudre est celui de l'évaluation de la probabilité p d'un événement sachant qu'il s'est produit k fois lors d'une série de t essais, le nombre théorique p étant défini par

$$p = \lim_{t \rightarrow \infty} \frac{k}{t}$$

Il est bien connu que plus le nombre d'essais t est grand, plus le rapport $\frac{k}{t}$ tendra à se rapprocher de p , conformément à la formule de *Bernoulli*

$$W_k^t(p) = \frac{t!}{k!(t-k)!} \cdot p^k \cdot (1-p)^{t-k} \quad (1)$$

qui nous donne la probabilité $W_k^t(p)$ que l'événement se produise k fois lors d'une série de t essais.

Diverses méthodes nous permettent d'évaluer le domaine dans lequel se situe p , en fonction de k et t , ces méthodes étant approchées et d'autant plus inexactes que t est petit. Dans la pratique, on ne dispose pas toujours d'un nombre d'essais t suffisamment élevé pour que ces méthodes puissent être utilisées et il importe pourtant de pouvoir déterminer avec précision le domaine dans lequel se situe p . Ainsi, par exemple, lorsqu'il s'agit d'évaluer la probabilité p qu'une nouvelle méthode opératoire soit suivie d'une issue mortelle, on ne peut sans autre augmenter le nombre d'essais t .

La méthode exacte de *Clopper* et *Pearson* (voir par exemple: *Mathematische Statistik* de B.L. *van der Waerden*, 1957, page 34) nous paraît entachée d'une erreur comme nous le verrons plus loin.

Nous allons énoncer et démontrer une règle I simple permettant de déterminer la probabilité d'exactitude $P(p)$ que la probabilité p d'un événement qui s'est produit k fois dans une série de t essais, soit renfermée dans le domaine $D(t, k)$ lorsqu'un tel domaine $D(t, k)$, $k = 0, 1, 2, \dots, t$, a été attribué à priori à chacun des k possibles.

La probabilité d'exactitude $P(p)$, fonction de la probabilité de l'événement, a un sens expérimental clair. En effet, supposons que N statisticiens effectuent chacun une série de t essais avec un même événement de probabilité inconnue p ; ils obtiendront en général l'une des valeurs $k = 0, 1, 2, \dots, t$ et admettront en conséquence que la probabilité p est renfermée dans le domaine correspondant $D(t, 0)$, $D(t, 1)$, $D(t, 2)$, ... ou $D(t, t)$. De ces N statisticiens, un certain nombre J obtiendront des domaines $D(t, k)$ qui renferment la probabilité p de l'événement expérimenté, tandis que

les $N - J$ autres obtiendront des domaines $D(t, k)$ ne renfermant pas la valeur de p . Par définition, nous posons:

$$P(p) = \lim_{N \rightarrow \infty} \frac{J}{N}$$

$P(p)$ est donc une fonction de la probabilité inconnue p et comme la valeur de p n'est pas connue, celle de P ne l'est pas non plus. Toutefois, la valeur minimum $P(p_{\min})$ nous permettra de déterminer la ou les valeurs p_{\min} de p pour lesquelles la fonction $P(p)$ prend la valeur minimum, ce qui constituera une indication intéressante.

Jusqu'à maintenant, nous avons considéré des domaines $D(t, k)$ donnés à priori de façon absolument arbitraire; nous établirons une règle II permettant de déterminer l'ensemble de domaines $D(t, k)$ assurant, dans l'ordre, les conditions suivantes: Une valeur minimum $P(p_{\min})$ donnée à priori, une somme des longueurs des domaines $D(t, k)$ qui soit minimum et une valeur de $\int_0^1 \sum_{k=0}^t D(t, k) W_k^t(p) dp$ qui soit maximum, les fonctions sous le signe \sum étant définies plus loin.

Règle I

Lorsqu'on effectue des séries de t essais avec un événement dont la probabilité est p , l'événement pouvant se produire $k = 0, 1, 2, \dots, t$ fois au cours d'une série d'essais, et que l'on fait correspondre à chacun des k possibles un domaine $D(t, k)$ quelconque situé entre 0 et 1, la probabilité $P(p)$ qu'un domaine $D(t, k)$, obtenu à la suite d'une série de t essais dans laquelle l'événement s'est produit k fois, renferme la valeur p de l'événement, est donnée par:

$$P(p) = \sum_{k=0}^t D(t, k) W_k^t(p) \quad (2)$$

expression dans laquelle $D(t, k) W_k^t(p)$ représente une fonction qui est égale à $W_k^t(p)$ dans le domaine $D(t, k)$ et qui est nulle en dehors de ce domaine.

Cette règle I est évidente: La probabilité d'obtenir k fois l'événement dans une série de t essais est égale à $W_k^t(p)$ et la probabilité d'obtenir l'événement k_p ou k_q ou k_r ou ... fois est évidemment égale à $W_{k_p}^t(p) + W_{k_q}^t(p) + W_{k_r}^t(p) + \dots$ Si maintenant, pour chacune des valeurs de p , $0 \leq p \leq 1$, nous limitons la somme précédente aux seules valeurs k_i pour lesquelles le domaine correspondant $D(t, k_i)$ renferme p , nous obtiendrons la probabilité d'exactitude $P(p)$ qu'un statisticien a d'obtenir un domaine $D(t, k)$

renfermant la probabilité p de l'événement, à la suite d'une série de t essais dans laquelle l'événement s'est produit k fois.

Règle II

L'ensemble des domaines $D(t,k)$, $k = 0,1,2\dots t$, assurant dans l'ordre les conditions suivantes:

- a) une valeur de $P(p_{\min})$ donnée;
- b) une somme des longueurs des domaines $D(t,k)$ qui soit minimum;
- c) une valeur de $\int_0^1 \sum_{k=0}^t D(t,k) W_k^t(p) dp$ qui soit maximum;

est obtenu en ordonnant, pour chacune des valeurs de p , $0 \leq p \leq 1$, les fonctions $W_k^t(p)$ par ordre de leurs valeurs décroissantes, et en n'en retenant, dans cet ordre, que le nombre minimum nécessaire pour que la somme de leurs valeurs en ce point p ne soit pas inférieure à $P(p_{\min})$, ce point p devant alors être renfermé dans les domaines $D(t,k)$ des fonctions $W_k^t(p)$ qui ont été retenues et ne devant pas être renfermé dans les domaines $D(t,k)$ des fonctions $W_k^t(p)$ qui n'ont pas été retenues.

Il est clair qu'en prenant, en chaque point p , le minimum de fonctions nécessaire pour que la somme des valeurs de ces fonctions au point p ne soit pas inférieure à la valeur prédéterminée $P(p_{\min})$ et en choisissant ces fonctions $W_k^t(p)$ par ordre de leurs valeurs décroissantes au point p , on satisfait automatiquement aux trois conditions a), b) et c) ci-dessus, car la condition a) ayant été satisfaite en prenant d'abord les fonctions ayant les valeurs les plus élevées, on en retiendra le plus petit nombre qui soit possible et, par conséquent, la somme des longueurs des domaines sera la plus petite possible, ce qui satisfera à la condition b); de plus, comme on a pris les fonctions $W_k^t(p)$ ayant les valeurs les plus élevées, la valeur de la fonction $P(p)$ sera maximum en chacun des points p , ce qui satisfera à la condition c).

On peut procéder graphiquement en traçant les différentes fonctions $W_k^t(p)$ comme représenté à la figure 1. En traçant une parallèle à l'axe des ordonnées, passant par l'abscisse p , on obtiendra les valeurs des fonctions $W_k^t(p)$ à leur point d'intersection avec cette parallèle, ces valeurs étant ordonnées par valeurs décroissantes de 1 à 0 sur cette parallèle. En traçant une parallèle à l'axe des abscisses qui coupe l'axe des ordonnées au point choisi $P(p_{\min})$, il suffira d'ajouter sur la parallèle à l'axe des ordonnées, successivement et par ordre des valeurs décroissantes, les ordonnées de ces différents points d'intersection jusqu'à ce qu'on atteigne la parallèle à l'axe des abscisses. Les seules fonctions $D(t,k) W_k^t(p)$ qui ne sont pas nulles en ce point p sont celles qu'il a été nécessaire d'utiliser pour atteindre la parallèle à l'axe des abscisses.

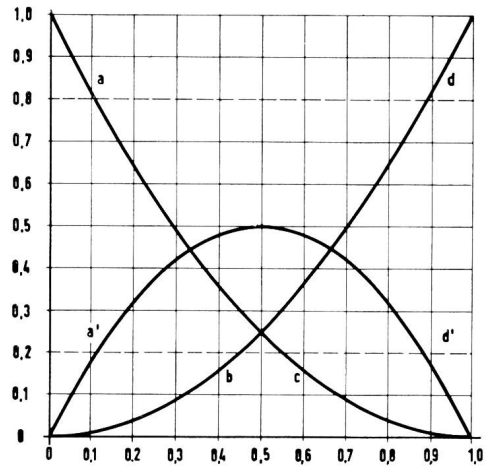


Fig. 1

Exemple

Nous allons considérer le cas simple: $t = 2$; $k = 0,1,2$.

La figure 1 représente les trois fonctions $W_0^2(p)$, $W_1^2(p)$ et $W_2^2(p)$.

Les figures 2a à 8a représentent différentes fonctions $P(p)$ possibles assurant des valeurs de $P(p_{\min})$ égales à 0,8; 0,75; 0,5 et 0,44.

La figure 9 représente la fonction $P(p)$ correspondant à un ensemble de domaines déterminé par la méthode de Clopper et Pearson, lorsque la probabilité d'exactitude garantie par cette méthode, et que nous désignerons aussi par $P(p_{\min})$ dans ce qui suit, est supérieure à 0,5. Comme on le verra plus loin, cette fonction $P(p)$ conduit, en appliquant la règle I, à une valeur de $P(p_{\min})$ égale à 0,41, donc inférieure à celle de 0,5 garantie par la méthode de Clopper et Pearson.

Les domaines $D(2,0)$, $D(2,1)$ et $D(2,2)$ ayant conduit aux fonctions $P(p)$ illustrées sur les figures 2a à 8a sont représentés par les trois segments supérieurs des figures correspondantes 2b à 8b, le domaine $D(2,0)$ correspondant au plus bas de ces trois segments, le domaine $D(2,1)$ au segment intermédiaire et le domaine $D(2,2)$ au segment supérieur.

Les trois segments inférieurs représentent dans le même ordre les domaines $D(2,0)$, $D(2,1)$ et $D(2,2)$, tels qu'ils sont déterminés par la méthode de Clopper et Pearson, lorsqu'on veut obtenir la même valeur de $P(p_{\min})$ que pour les trois segments supérieurs.

En traçant sur la figure 1 des droites horizontales d'ordonnées 0,8 et 0,2, on obtient les points d'intersection a, d, respectivement a', b, c et d', qui déterminent les limites des différents domaines possibles.

En appliquant intégralement la règle II, on obtient la fonction $P(p)$ représentée à la figure 2a; mais on pourrait aussi obtenir la même $P(p_{\min}) = 0,8$ en ne tenant pas compte de la condition c) de la règle II. Une fonction $P(p)$ ainsi obtenue est représentée à la figure 3a. La longueur totale des domaines (voir

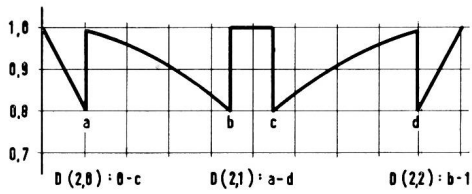


Fig. 2a $P(p_{\min}) = 0,8$ Fig. 2b

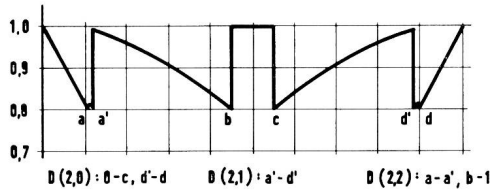


Fig. 3a $P(p_{\min}) = 0,8$ Fig. 3b

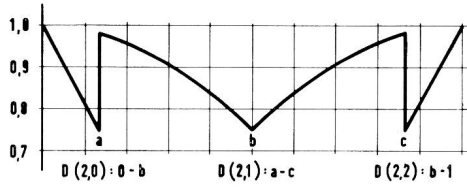
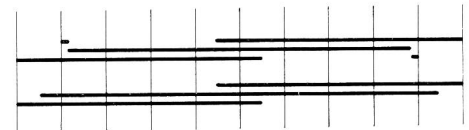


Fig. 4a $P(p_{\min}) = 0,75$ Fig. 4b

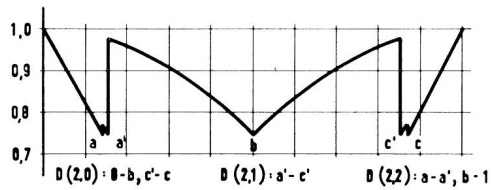
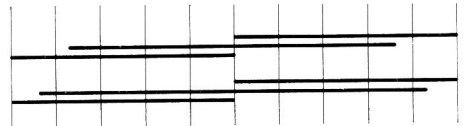


Fig. 5a $P(p_{\min}) = 0,75$ Fig. 5b

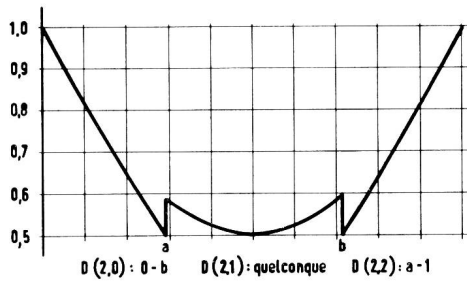
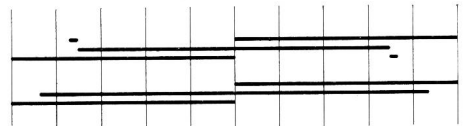


Fig. 6a $P(p_{\min}) = 0,5$ Fig. 6b

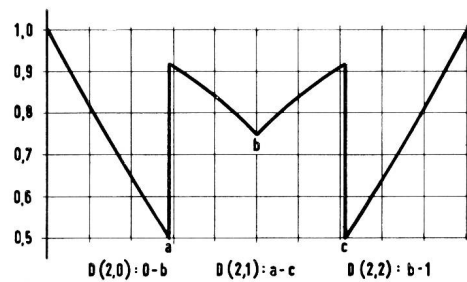
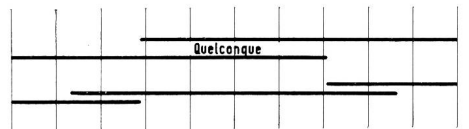


Fig. 7a $P(p_{\min}) = 0,5$ Fig. 7b

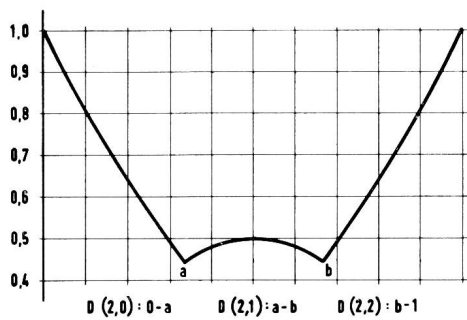
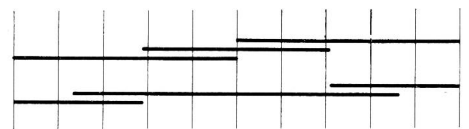
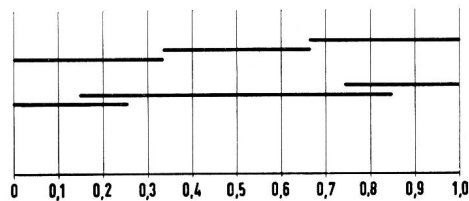


Fig. 8a $P(p_{\min}) = 0,44$ Fig. 8b



figures 2b et 3b) est la même dans les deux cas, mais $\int_0^1 P(p)dp$ est évidemment plus grande dans le premier cas que dans le second et il est clair que le choix des domaines correspondant aux figures 2a et 2b est plus judicieux que celui correspondant aux figures 3a et 3b, car si la probabilité p était renfermée entre les points a et a' ou d' et d, la probabilité d'exactitude $P(p)$ serait plus élevée dans le premier cas que dans le second.

Les figures 4a et 5a respectivement 4b et 5b représentent de même deux choix possibles, lorsque $P(p_{\min})$ a été fixée à 0,75.

Les figures 6a et 7a représentent deux choix possibles des domaines, dans le cas où $P(p_{\min})$ a été fixée à 0,5.

Le cas de la figure 6a et de la figure 6b mérite une attention particulière; il montre que le domaine $D(2,1)$ peut être choisi de façon tout à fait arbitraire et être de longueur nulle, par exemple.

Cela peut paraître paradoxal; en réalité, ce résultat correspond bien à notre définition de la probabilité d'exactitude $P(p)$. Supposons que la probabilité p de l'événement se situe au voisinage des points a, 0,5 ou b, les plus défavorables (voir figure 6a). Si l'on admet que les statisticiens obtenant $k = 1$ se trompent tous dans l'évaluation du domaine $D(2,1)$, puisqu'ils peuvent choisir un domaine quelconque, par exemple constitué par un point unique, il y aura 50% de statisticiens dans l'erreur et 50% qui attribueront un domaine renfermant effectivement la valeur de p .

Il est clair qu'un tel choix des domaines $D(2,k)$ n'a aucun sens, puisque les statisticiens ayant obtenu $k = 1$ seraient autorisés à faire un choix certainement erroné.

Cela explique la nécessité de la condition c) de la règle II puisque dans le cas de la figure 6a, $P(p_{\min}) = 0,5$ et la somme des longueurs des domaines est minimum, les conditions a) et b) étant donc satisfaites. La figure 7a représente le cas où la condition c) a été satisfaite en plus des conditions a) et b).

On peut voir aux figures 7a et 7b, d'une part, et aux figures 8a et 8b, d'autre part, qu'il suffit de passer d'une valeur de $P(p_{\min})$ égale à 0,5 à une valeur égale à 0,44, pour que la somme des longueurs des domaines $D(t,k)$ diminue d'environ 0,4. Pour cette valeur de 0,44 il n'y a qu'un choix possible des domaines qui satisfait aux conditions a) et b).

La règle de Clopper et Pearson

Cette règle peut s'énoncer ainsi:

A. Si dans une série de t essais, un événement s'étant produit k fois, on considère la valeur de la probabilité p de cet événement comme étant comprise entre 0 et p_1^k , avec p_1^k définie par l'équation

$$\sum_{i=0}^k W_i^t(p_1^k) = \beta \quad (3)$$

la probabilité que l'on commette une erreur, c'est-à-dire la probabilité que $p_1^k < p < 1$, est inférieure à β .

B. Si, à la suite de ce résultat, on considère que p est située entre p_2^k et 1, avec p_2^k définie par l'équation

$$\sum_{i=k}^t W_i^t(p_2^k) = \beta \quad (4)$$

la probabilité que l'on commette une erreur est aussi inférieure à β et si l'on considère p comme étant comprise entre p_1^k et p_2^k , la probabilité de commettre une erreur est inférieure à 2β .

Nous verrons que, si la partie A de la règle de Clopper et Pearson correspond aux résultats prévus par la règle I, il n'en est pas de même de la partie B faisant intervenir les deux limites p_1^k et p_2^k . En effet, dire que l'on commet une erreur inférieure à β , lorsqu'on suppose p comprise entre 0 et p_1^k , signifie que la probabilité d'exactitude est supérieure à $1 - \beta$. Or, il est bien évident qu'en appliquant la règle I au cas de domaines $D(t,k)$ compris entre 0 et des valeurs p_1^k définies par les équations

$$S_k(p_1^k) = \sum_{i=0}^k W_i^t(p_1^k) = \beta \quad (5)$$

la fonction

$$P(p) = \sum_{i=0}^t D(t,i)W_i^t(p)$$

prendra en chacun des t points p_1^k ($k = 0,1,2,\dots, t-1$) la valeur $1 - \beta$ et restera supérieure à cette valeur entre les points p_1^k puisque d'une part les t fonctions $S_k(p)$ vont en décroissant de 1 à 0 lorsque p varie de 0 à 1 et que d'autre part $\sum_{i=0}^t W_i^t(p) = 1$. En conséquence, la valeur $P(p_{\min})$ prise par la fonction $P(p)$ sera bien égale à $1 - \beta$.

Nous allons maintenant, en nous reportant à l'exemple décrit plus haut pour $t = 2$, déterminer les limites des différents domaines $D(2,k)$ par la méthode de Clopper et Pearson de façon que la probabilité d'exactitude garantie par cette méthode et que nous désignerons aussi par $P(p_{\min})$, soit égale successivement à 0,8; 0,75; 0,5; 0,44 comme nous l'avons fait avec la règle II. Il nous suffira pour cela de résoudre les équations suivantes en faisant $P(p_{\min})$ égale successivement à 0,8; 0,75; 0,5 et 0,44.

$$\begin{aligned} W_0^2(p_1^0) &= 1 - P(p_{\min}) \\ W_0^2(p_1^1) + W_1^2(p_1^1) &= \frac{1}{2} [1 - P(p_{\min})] \\ W_1^2(p_2^1) + W_2^2(p_2^1) &= \frac{1}{2} [1 - P(p_{\min})] \\ W_2^2(p_2^2) &= 1 - P(p_{\min}) \end{aligned} \quad (6)$$

Cela nous détermine les trois domaines $D(2,0) = 0 - p_1^0$, $D(2,1) = p_2^1 - p_1^1$, et $D(2,2) = p_2^2 - 1$.

Ces trois domaines, pour chacun des $P(p_{\min}) = 0,8, 0,75, 0,5$ et 0,44, sont représentés aux figures 2b à 8b, par les trois segments inférieurs.

Nous voyons que ces domaines ne correspondent nullement à ceux que nous avons déterminés par la règle II.

Les auteurs mentionnés ne donnent pas de définition de la probabilité β ; toutefois, comme la partie A de la règle de Clopper et Pearson conduit au même résultat que la règle I, nous pourrions admettre que β a la même signification que $P(p_{\min})$, à cette différence près que β représente la plus grande probabilité d'erreur et $P(p_{\min})$ la plus petite probabilité d'exactitude,

$$P(p_{\min}) = 1 - \beta.$$

Or il est évident, d'après la démonstration que nous avons faite de la règle I, que lorsqu'un ensemble de domaines $D(t, k)$ a été choisi, lequel conduit à une certaine valeur de $P(p_{\min})$, on ne peut sans autre lui superposer un second ensemble, même si ce dernier considéré seul conduit à la même valeur de $P(p_{\min})$; c'est pourtant ce qui a été fait dans la partie B de la règle de Clopper et Pearson. Dans cette partie B, ces auteurs ne considèrent que les deux limites p_1^k et p_2^k se rapportant à un seul k , et soustraient de 1 la probabilité maximum β qu'à p d'être comprise dans le domaine $p_1^k - 1$ et la probabilité maximum β qu'à p d'être comprise dans le domaine $0 - p_2^k$ pour obtenir la probabilité minimum $1 - 2\beta$ qu'à p d'être comprise entre p_2^k et p_1^k . Or, cela est manifestement erroné, car lorsqu'on détermine une probabilité, il faut tenir compte de tous les résultats possibles et dans le cas ci-dessus des autres valeurs possibles de k , lesquelles conduisent à d'autres limites p_1^k et p_2^k .

Nous allons discuter maintenant un exemple numérique précis où la méthode de Clopper et Pearson conduit à des résultats inexacts. Supposons une série de deux essais, $t = 2$, dans laquelle l'événement peut se produire 0, 1 ou 2 fois. Nous allons, pour chacune de ces trois possibilités, déterminer les domaines auxquels nous conduit la méthode de Clopper et Pearson, lorsque la probabilité d'exactitude doit être égale à 0,5. En faisant $P(p_{\min}) = 0,5$ dans les équations 6), nous obtenons les résultats suivants: $p_1^0 = 0,293$; $p_1^1 = 0,865$; $p_2^1 = 0,135$ et $p_2^2 = 0,707$.

Si maintenant nous déterminons la fonction $P(p)$ correspondant aux domaines ainsi calculés, conformément à la règle I, nous obtenons la courbe représentée à la figure 9 où nous voyons que $P(p_{\min})$ est égale à 0,41 et non à 0,5 comme déterminé par la méthode de Clopper et Pearson. Nous voyons également sur cette figure 9 que la fonction $P(p)$ prend les valeurs minimums $P(p_{\min})$ pour $p_{\min} = 0,29$ et $p_{\min} = 0,71$.

Imaginons des statisticiens qui doivent déterminer expérimentalement une probabilité p dont la valeur est voisine de l'une de ces deux valeurs de p_{\min} , par exemple $1/3$, en faisant une série de deux essais et de façon à obtenir une probabilité d'exactitude $P(p_{\min}) = 0,5$.

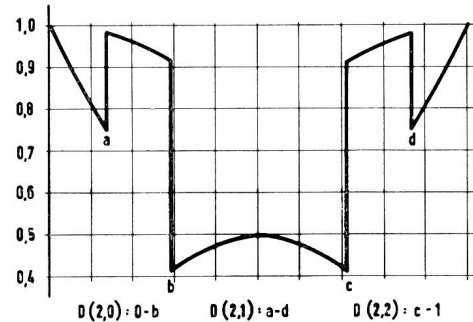


Fig. 9

Un tel événement peut par exemple être obtenu par un dé à six faces dont deux seulement sont marquées; la probabilité de sortir une face marquée sera évidemment égale à $2 \times 1/6 = 1/3$. En faisant $p = 1/3$ dans la formule de Bernoulli, nous voyons que, de tous ces statisticiens:

$$\frac{2!}{0! 2!} (1/3)^0 (1-1/3)^2 = 4/9$$

obtiendront $k = 0$, situeront p entre 0 et 0,293 et seront donc dans l'erreur;

$$\frac{2!}{1! 1!} (1/3) (1-1/3) = 4/9$$

obtiendront $k = 1$, situeront p entre 0,135 et 0,865 et seront dans la vérité;

$$\frac{2!}{2! 0!} (1/3)^2 (1-1/3)^0 = 1/9$$

obtiendront $k = 2$, situeront p entre 0,707 et 1 et seront donc dans l'erreur.

Nous voyons donc que 44,4% des statisticiens obtiendront, dans ce cas évidemment défavorable, un résultat exact et non 50% comme le garantit la règle de Clopper et Pearson.

En adoptant, au contraire, l'ensemble des domaines défini par la règle II (voir figures 7a et 7b), les statisticiens obtenant $k = 0$ et $k = 1$ choisiront un domaine renfermant $p = 1/3$ et par conséquent 88,8% de ceux-ci obtiendront un résultat exact au lieu de 44,4% seulement.

Ces valeurs de 44,4% et 88,8% peuvent être lues directement aux figures 9 respectivement 7a et correspondent à la valeur prise par la courbe portée en ordonnée lorsque l'abscisse est égale à 0,33.

Adresse de l'auteur: René Wohlers, ingénieur, Landhausweg 34, Berne.