# Problème de deux races

**R. de Misès** (Istanbul)

Voilà un problème qui se pose souvent dans la statistique générale ou dans la statistique biologique, ou bien dans la science d'actuaires:

Un ensemble comprenant $n$ individus est composé de deux classes, $\alpha n$ individus appartenant à une classe $A$ et $\beta n = n - \alpha n$ à une classe $B$. Soit $z$ un caractère distinctif quelconque, disons par exemple la longueur d'un individu de l'ensemble. On suppose l'existence, pour chacune des deux classes, d'une distribution des probabilités $A(z)$ respectivement $B(z)$. Notamment $A(z)$ signifie* la probabilité, pour que le caractère distinctif d'un individu de la classe $A$ ne surpasse pas la valeur $z$. On demande: quelle est la probabilité $P(x)$ pour que, parmi les $m$ individus présentant les plus grandes valeurs de $z$ dans l'ensemble, se trouvent $x$ individus de la classe $A$ et $y = m - x$ de la classe $B$?

Il est évident qu'une telle question est sensée même s'il n'y a pas moyen à mesurer d'une façon absolue les valeurs de $z$. Il suffit qu'on sache ranger les individus d'après les valeurs de $z$, de sorte que le cas $z_1 > z_2$ est nettement distinct du cas $z_1 < z_2$. Encore peut on appliquer les résultats si la statistique ne donne qu'une séparation des individus en deux catégories, telles que les valeurs de $z$ dans la première catégorie sont supérieures à ceux de la deuxième. Par exemple on sera disposé à admettre que, parmi les physiciens d'un pays ceux qui sont titulaires du prix Nobel soient „les plus grands". Alors, si les habitants de ce pays appartiennent à deux races différentes l'une de l'autre, notre recherche y sera applicable.

En ce qui suit nous donnons tout d'abord ($n^0$ 1) l'expression générale pour $P(x)$ sous forme d'une intégrale de Stieltjes (6) qui se ramène, en supposant $A(z)$ et $B(z)$ continues, à une intégrale définie au sens ordinaire. Puis ($n^0$ 2) nous en déduisons la formule (8), d'ailleurs très simple et facile à comprendre, pour le cas présentant un intérêt spécial, où les deux probabilités $A$ et $B$ sont égales entre elles; l'étude détaillée de la distribution $P(x)$ n'y présente aucune difficulté. Pour utiliser l'expression générale (6) il faut passer à la limite pour $n$ infini, ce qui nous fournira des formules approchées, applicables aux cas où $n$ est un entier assez grand. Correspondant aux deux façons classiques de passage à la limite, dues à Poisson et à Laplace, il y a dans notre problème aussi deux manières différentes à envisager les valeurs infiniment grandes de $n$. Nous traitons en premier lieu ($n^0$ 3) le cas où $m$ reste fini, $n$ tendant vers l'infini, c'est-à-dire où on ne considère qu'un nombre restreint de plus grandes valeurs de $z$ dans un ensemble de très grande étendue. Le deuxième cas ($n^0$ 4) est caractérisé par le fait que le rapport

*

$m:n$ reste fixe, de sorte que l'objet d'observation est formé par une partie considérable de l'ensemble total. Du point de vue de l'analyse mathématique il s'agit dans tous les deux cas de certaines généralisations de procédés connus d'intégration asymptotique.

On s'intéressera aussi pour le problème inverse: étant données les valeurs de $x$ et $y$, quelle est la probabilité qu'elles résultent d'une certaine relation entre $A(z)$ et $B(z)$? Nous donnons (n⁰ 5) quelques indications concernant cette question, sans l'épuiser au fond. Enfin sont signalées (n⁰ 6) deux applications, l'une d'une façon générale, l'autre en entrant dans le calcul détaillé, exemple qui, par son sujet même, prêtera peut-être un intérêt spécial.

## 1. Mise en équation du problème

Sont donnés deux collectifs primordiaux $C_A$ et $C_B$, à une dimension [1], avec leurs distributions $A(z)$ et $B(z)$. On sait que les fonctions $A(z)$ et $B(z)$ sont monotones, non-décroissantes et que

$$A(-\infty)=B(-\infty)=0, \quad A(\infty)=B(\infty)=1. \tag{1}$$

Nous supposons de plus qu'elles soient continues, de sorte que des probabilités „concentrées" restent exclues. Les deux équations

$$x=A(z) \quad \text{et} \quad y=B(z)$$

définissent une courbe à ordonnées non-décroissantes, menant du point $x=$ $=y=0$ au point $x=y=1$. On peut donc considérer $B$ comme une fonction monotone de $A$, en précisant d'une manière convenable les valeurs de $B$ dans les points où $A(z)$ reste constant. En tout cas $B(A)$ est intégrable et l'intégrale ne dépend pas des valeurs arbitraires, attribuées à $B$ dans les points dont nous venons de parler.

Des deux collectifs donnés $C_A$ et $C_B$ on passe par „composition" à un collectif $C$ à $n$ dimensions. Un élément de $C$ est composé de $a=\alpha n$ éléments de $C_A$ et de $b=\beta n=n-a$ éléments de $C_B$, son caractère distinctif se compose de $a$ valeurs $z_1', z_2', \ldots, z_a'$ et de $b$ valeurs $z_1'', z_2'', \ldots, z_b''$. La distribution dans $C$ est donnée par la règle bienconnue de multiplication des probabilités. Soient $I_k'$, $I_\lambda''$ des intervalles quelconques (à une dimension) et $\Delta_k A$ respectivement $\Delta_\lambda B$ les probabilités correspondant à ces intervalles dans $C_A$ respectivement $C_B$, alors le produit

$$\Delta_1 A \cdot \Delta_2 A \ \ldots \ \Delta_a A \cdot \Delta_1 B \cdot \Delta_2 B \ \ldots \ \Delta_b B \tag{2}$$

détermine la probabilité dans $C$, pour que $z_k'$ tombe dans $I_k'$ et $z_\lambda''$ dans $I_\lambda''$ $(k=1, 2, \ldots, a, \ \lambda=1, 2, \ldots, b)$.

Par sommation (intégration d'après Stieltjes) des expressions (2) on trouve la probabilité pour un domaine quelconque de l'espace caractéristique de $C$. L'intégrale existe surtout pour des domaines dont le contour n'est formé que

[1] Je me sers ici et dans ce qui suit de certaines expressions introduites dans mon „Cours des probabilités" (1931) et dans mes mémoires antérieurs; mais le lecteur, familiarisé avec les éléments de la théorie classique, ne trouvera pas des difficultés, même sans recourir à ces sources.

par de hyperplans, c'est-à-dire par d'ensembles de points, définis par un nombre quelconque d'équations linéaires. On déduit immédiatement de (2) que la probabilité s'annulle pour chaque ensemble situé sur un hyperplan même. Il en suit qu'on puisse supprimer de prime abord tous les points, pour lesquels deux ou plusieurs coordonnées sont égales entre elles.

Soit donc $z_k'$, $z_\lambda''$ un système de $n$ valeurs différentes l'une de l'autre. Pour un tel système on peut toujours discerner s'il jouisse ou non de la propriété suivante: parmi les $m$ plus grandes valeurs ($m \leqslant a$, $m \leqslant b$) se trouvent $x$ valeurs de la première catégorie (des $z_k'$) et $m - x = y$ valeurs de la deuxième (des $z_\lambda''$). L'ensemble des points satisfaisant à cette condition constitue le domaine $E$. La probabilité cherchée $P(x)$ est l'intégrale de Stieltjes de l'expression (2), étendue au domaine $E$.

On peut décomposer $E$ en deux parties $E_1$ et $E_2$ où $E_1$ est caractérisé par le fait que la plus petite parmi les $m$ plus grandes valeurs appartient aux $z_k'$, tandis que dans $E_2$ la plus petite des $m$ valeurs envisagées est un $z_\lambda''$. Cherchons maintenant à évaluer l'intégrale de (2) pour le domaine $E_1$, qui, d'ailleurs, n'est limité que de hyperplans de la forme $z_\rho' = z_\sigma'$ ou $z_\rho' = z_\sigma''$ etc.

Pour chaque point de l'ensemble $E_1$ il existe un intervalle $z < z' \leqslant z + h$ ($h > 0$) qui joue le rôle suivant: 1) une seule valeur des $z'$ appartient à l'intervalle, 2) pour $(x-1)$ valeurs de la première catégorie on a $z' > z + h$ et pour $(a - x)$ de ces valeurs $z' \leqslant z$, 3) pour $y$ valeurs de la deuxième catégorie on a l'inégalité $z'' > z + h$ et pour les autres $(b - y)$ l'inégalité $z'' \leqslant z$. La probabilité pour qu'une telle répartition des $n$ valeurs soit réalisée par un certain arrangement de coordonnées, par exemple par

$$z < z_1' \leqslant z + h,\ z_2' > z + h,\ z_3' > z + h,\ \ldots,\ z_x' > z + h,\ z_{x+1} < z,\ \ldots \text{ etc.},$$

est d'après (2) égale à

$$[A(z)]^{a-x} \cdot \Delta A \cdot [1 - A(z+h)]^{x-1} \cdot [B(z)]^{b-y} [1 - B(z)]^y, \tag{3}$$

où $\Delta A$ signifie $A(z+h) - A(z)$. Le nombre total des différents arrangements réalisant la même répartition des $a + b = n$ valeurs se trouve d'après les règles de l'analyse combinatoire égal à

$$\frac{a!}{(a-x)!\ 1!\ (x-1)!} \cdot \frac{b!}{(b-y)!\ y!} = x \binom{a}{x} \binom{b}{y}. \tag{4}$$

Désignons par $Q_h(z)$ le produit des deux expressions (3) et (4). La somme infinie $Q_h(z) + Q_h(z-h) + Q_h(z+h) + Q_h(z+2h) + Q_h(z-2h) + \ldots$ représente la probabilité correspondant à une certaine partie de $E_1$, mais elle n'épuise pas tout l'ensemble $E_1$, puisque les points $z_k'$, $z_\lambda''$ mis en compte sont soumis à la restriction suivante: la plus petite des $m$ plus grandes valeurs parmi les $z_k'$, $z_\lambda''$ est „isolée", de sorte qu'un intervalle de l'étendue $h$ autour d'elle reste libre de toute autre valeur. On se débarasse de cette restriction, étrangère à la définition de $E_1$, en diminuant $h$ vers zéro, ou en passant à la limite $h = 0$ dans la somme envisagée. Cette limite est, par définition, l'intégrale de Stieltjes, étendue de $z = -\infty$ à $z = \infty$:

$$\int x \binom{a}{x} \binom{b}{y} [A(z)]^{a-x} [1 - A(z)]^{x-1} [B(z)]^{b-y} [1\ B(\mid z)]^y \, dA(z). \tag{5}$$

Comme il était déjà établi que $B$ est une fonction intégrable de $A$, on peut considérer (5) comme une intégrale définie avec la variable d'intégration $A$, étendue à l'intervalle de $A=0$ à $A=1$.

La deuxième partie $E_2$ de $E$ livrant la même expression avec la substitution $B$ pour $A$, $b$ pour $a$ et $y$ pour $x$ et vice-versa on a finalement

$$P(x)=\binom{a}{x}\binom{b}{y}[x\int_0^1 A^{a-x}(1-A)^{x-1}B^{b-y}(1-B)^y\,dA+$$

$$+y\int_0^1 A^{a-x}(1-A^x)B^{b-y}(1-B)^{y-1}\,dB]. \tag{6}$$

On y voit que $P(x)$, les $a, b, x, y$ donnés, ne dépend que de la relation entre $A$ et $B$; la probabilité $P(x)$ ne change pas, si l'on transforme la variable indépendante $z$ de la façon $z=\varphi(\bar{z})$ où $\varphi$ est une fonction monotone, croissante et continue. La formule (6) renferme la solution générale du problème.

## 2. Cas de deux distributions égales $A=B$

L'équation (6) se simplifie foncièrement au cas que les deux distributions $A(z)$ respectivement $B(z)$ dans les collectifs initiaux sont égales entre elles. Les deux intégrales au deuxième membre de (6) ne diffèrent pas l'une de l'autre et, puisque $a+b=n$, $x+y=m$, il vient:

$$\overline{P}(x)=\binom{a}{x}\binom{b}{y}m\int_0^1 A^{n-m}(1-A)^{m-1}\,dA=\frac{\binom{a}{x}\binom{b}{y}}{\binom{n}{m}} \tag{7}$$

ou, en prenant $y=m-x$,

$$P(x)=\frac{m!\,(n-m)!\,a!\,b!}{n!x!\,(a-x)!\,(m-x)!\,(b-m+x)!}. \tag{8}$$

Le numérateur au dernier membre de (7) est le coefficient de $t^m u^x$ dans le développement du produit

$$(1+tu)^a(1+t)^b.$$

Il en suit que la „fonction génératrice" de la distribution $P(x)$ est donnée par

$$\overline{P}(u)=\sum_x P(x)u^x=\frac{(n-m)!}{n!}\frac{d^m}{dt^m}\{(1+tu)^a(1+t)^b\}_{t=0}. \tag{9}$$

En introduisant $u=1$ on en déduit facilement que la somme des $P(x)$ est égale à l'unité, cette somme, celle de (9) et les suivantes, étant étendue aux valeurs de $x=0$ à $x=m$. La première dérivée de (9) nous donne

$$\sum_x x\,P(x)\,u^{x-1}=a\frac{(n-m)!}{n!}\frac{d^m}{dt^m}\{t(1+tu)^{a-1}(1+t)^b\}_{t=0} \tag{10}$$

et, en introduisant $u=1$, on trouve pour la valeur moyenne ou l'espérance mathématique de $x$:

$$\sum_x x\,P(x)=a\frac{(n-m)!}{n!}m!\binom{n-1}{m-1}=\frac{am}{n}=\alpha m. \tag{11}$$

Par conséquent, la valeur moyenne de $y = m - x$ est

$$m - \alpha m = \beta m.$$

Il n'y a rien d'étonnant dans ces résultats: si l'on suppose les deux probabilités initiales $A$ et $B$ égales entre elles, il faut attendre qu'en moyen, les nombres $x$ et $y$ des individus des deux classes, parmi les $m$ individus aux plus grandes valeurs de $z$, soient en proportion $\alpha : \beta$, proportion des étendues mêmes des deux classes dans l'ensemble total.

La deuxième dérivée de (9) ou la première de (10), pour $u = 1$, mène à

$$\sum_x x(x-1)P(x) = a(a-1)\frac{(n-m)!}{n!}m!\binom{n-2}{m-2} = a(a-1)\frac{m(m-1)}{n(n-1)}. \quad (12)$$

De (12) on déduit par des simples calculs la déviation ou l'écar quadratique de la distribution (7):

$$\sum_x (x - \alpha m)^2 P(x) = \frac{abm(n-m)}{n^2(n-1)} = \alpha\beta m \frac{n-m}{n-1}. \quad (13)$$

On peut, d'ailleurs, étudier le cours de la fonction $P(x)$ en partant du quotient, déduit de (8):

$$\frac{P(x+1)}{P(x)} = \frac{a-x}{x+1}\,\frac{m-x}{b-m+x+1}. \quad (14)$$

En égalant le numérateur et le dénominateur on y voit que $P(x)$ est supérieur à $P(x-1)$ autant que

$$x < \frac{(m+1)(a+1)}{n+2} = x_0. \quad (15)$$

Donc, $P(x)$ est croissant de $x = 0$ jusqu'à $x = [x_0]$, puis décroissant. Pour $x$ entier les valeurs $P(x_0)$ et $P(x_0 - 1)$ sont égales entre elles et supérieures$_0$ à toutes autres. On éprouve facilement que sous les conditions données $m \leqslant a$, $m \leqslant b$

$$x_0 - 1 < m\frac{a}{n} = \alpha m < x_0. \quad (16)$$

Il en suit que la valeur moyenne $\alpha m$, si elle est un nombre entier, est en même temps la valeur la plus probable; dans l'autre cas le maximum de la probabilité est atteint au moins pour l'un des deux nombres entiers voisins de $\alpha m$.

Passons maintenant à la limite pour $n$ infini. Nous envisagerons deux sortes de limites: 1) en fixant $m$, $x$, $y$ pendant que $n$ tend vers $\infty$, 2) en augmentant $m$, $x$, $y$ en proportion avec $n$. En tout cas $\alpha$, $\beta$ sont des valeurs fixes.

Au premier cas, partant de (8), on tombe sur

$$\lim_{n\to\infty} P(x) = \frac{m!\,a^x b^{m-x}}{n^m x!\,(m-x)!} = \binom{m}{x}\alpha^x\beta^{m-x}. \quad (17)$$

Voilà une distribution, dite de Bernoulli, aux probabilités élémentaires $\alpha$, $\beta$. La probabilité en question $P(x)$ est donc, pour des $n$ assez grands, la même que la probabilité du fait que, dans une suite de $m$ épreuves répétées, l'événement, dont la probabilité est $\alpha$, se produit $x$ fois. La valeur moyenne de (17) est $\alpha m$, l'écart quadratique $\alpha\beta m$.

Pour le deuxième cas écrivons

$$x = \xi n, \quad y = \eta n, \quad m = \mu n = (\xi + \eta)\, n, \tag{18}$$

où $\mu$ est une constante donnée et $\xi$, $\eta$ sont des variables. En passant à $n$ infini, nous écrivons

$$\xi = \alpha\mu + \frac{u}{\sqrt{n}}, \qquad \eta = \beta\mu - \frac{u}{\sqrt{n}} \tag{19}$$

et considérons $u$ comme borné. Utilisant la formule de Stirling on trouve

$$\lim_{n\to\infty} \sqrt{n}\, P(\alpha m + u\sqrt{n}) = \frac{1}{\sqrt{2\pi\alpha\beta\mu\,(1-\mu)}}\, e^{-\frac{u^2}{2\alpha\beta\mu\,(1-\mu)}}. \tag{20}$$

Nous supprimons ici la déduction complète de (20) puisqu'elle sera donnée ci-dessous sous des conditions plus générales. La distribution gaussienne (20), pour $x$ comme variable indépendante, a la valeur moyenne $\alpha m$ et l'écart quadratique $\alpha\beta\mu\,(1-\mu)\, n$, ce qui est en concordance avec (13). Toutefois il faut signaler que le coefficient de $u^2$ dans (20) n'est pas celui qui corresponderait au cas de $m$ répétitions d'une épreuve aux probabilités élémentaires $\alpha$, $\beta$.

### 3. Cas général $A \neq B$. Limite pour $n$ infini, $m$ fini

Dans le cas général où les deux probabilités initiales $A$ et $B$ sont différentes l'une de l'autre, on parvient à des formules utilisables si l'on prétend que le nombre d'individus est assez grand, les rapports $a:n = \alpha$ et $b:n = \beta$ restant constants. Comme il était déjà signalé tout à l'heure, il y faut distinguer deux problèmes différents selon la grandeur de $m$. Si l'on n'envisage, dans un ensemble assez étendu, qu'un nombre restreint de quelques individus, dont la valeur du caractère distinctif est extrême, il faut regarder $x$, $y$ et $m$ comme nombres restant finis, pendant que $n$ tend vers l'infini. Au contraire, si l'objet d'observation est formé par une partie considérable de l'ensemble total, on regardera les rapports $m:n$, $x:n$ etc. comme restant finis, $n$ tendant vers l'infini. Nous verrons que, pour chacun des deux cas, des différentes propriétés des distributions $A(z)$, $B(z)$ joueront le róle décisif.

Pour le premier problème, $m$ fini, nous supposons que la fonction $B(A)$, signalée ci-dessus, soit continue et dérivable au point $A = 1$, $B = 1$, et qu'il soit de même pour $A(B)$. En autres termes, si $\lambda$ signifie une constante positive et $\eta(A)$ une fonction s'annulant pour $A = 1$ et continue pour des valeurs assez petites de $1 - A$:

$$\frac{1-B}{1-A} = \lambda\,[1 + \eta(A)], \quad \lim_{A\to 1} \frac{1-B}{1-A} = \lambda. \tag{21}$$

On en déduit immédiatement

$$\lim_{A\to 1} \frac{\ln(A^\alpha B^\beta)}{1-A} = \alpha \lim \frac{\ln A}{1-A} + \beta \lim \frac{1-B}{1-A}\,\frac{\ln B}{1-B} = -\alpha - \beta\lambda < 0. \tag{22}$$

Posons $\alpha + \beta\lambda = r$ et désignons par $\rho(A)$ une fonction jouissant de mêmes propriétés que $\eta(A)$; on obtient

$$\frac{\ln(A^\alpha B^\beta)}{1-A} = -r\,[1 + \rho(A)]. \tag{23}$$

Nous allons démontrer que la première intégrale dans l'expression (6)

$$I_A = \int_0^1 A^{\alpha n - x} B^{\beta n - y} (1 - A)^{x-1} (1 - B)^y \, dA, \tag{24}$$

où $\alpha$, $\beta$, $x$, $y$ restent constants, pendant que $n$ tend vers l'infini, satisfait à l'équation asymptotique

$$\lim_{n \to \infty} n^m I_A = (m-1)! \, r^{-m} \lambda^y, \quad m = x + y. \tag{25}$$

Soit $A_1$ une valeur quelconque $< 1$ et $I_1$ l'intégrale (24) étendue de 0 à $A_1$. Pour le reste nous remplaçons $A$ par une nouvelle variable $t$, moyennant

$$t = nr(1 - A), \quad t_1 = nr(1 - A_1) \tag{26}$$

de sorte que l'intégrale se transforme en

$$\int_{A_1}^1 (A^\alpha B^\beta)^n \cdot (1 - A)^{x-1} (1 - B)^y \cdot A^{-x} \cdot B^{-y} \, dA = (nr)^{-m} \lambda^y \int_0^{t_1} f(t) \, dt \tag{27}$$

où d'après (21) et (23)

$$f(t) = e^{-t(1+\rho)} \cdot t^{m-1} (1 + \eta)^y \cdot \left[1 - \frac{t}{nr}\right]^{-x} \left[1 - \frac{t\lambda(1+\eta)}{nr}\right]^{-y}. \tag{28}$$

Soit

$$\varphi(t) = f(t) - e^{-t} t^{m-1} \tag{29}$$

et citons la formule connue

$$\int_0^\infty e^{-t} t^{m-1} dt = (m-1)! \tag{30}$$

Pour une valeur intermédiaire $t_0 < t_1$ quelconque on a

$$\lambda^{-y} (nr)^m I_A = \lambda^{-y} (nr)^m I_1 + \int_0^{t_0} f(t) \, dt + \int_{t_0}^{t_1} f(t) \, dt = \lambda^{-y} (nr)^m I_1 +$$

$$+ \int_0^{t_0} \varphi(t) \, dt + \int_0^\infty e^{-t} t^{m-1} \, dt + \int_{t_0}^{t_1} f(t) \, dt - \int_{t_0}^\infty e^{-t} t^{m-1} \, dt \tag{31}$$

d'où on déduit, en utilisant (30):

$$|\lambda^{-y} (nr)^m I_A - (m-1)!| < \lambda^{-y} (nr)^m I_1 + \int_0^{t_0} |\varphi(t)| \, dt +$$

$$+ \int_{t_0}^{t_1} f(t) \, dt + \int_{t_0}^\infty e^{-t} t^{m-1} \, dt. \tag{32}$$

Une borne supérieure du premier terme au deuxième membre se trouve sous la forme:

$$(nr)^m I_1 < (nr)^m \int_0^A A_1^{\alpha n - x} \, dA < (nr)^m A_1^{\alpha n - x + 1}. \tag{33}$$

Cette expression s'annule, $n$ augmentant infiniment, pour chaque valeur $A_1 < 1$. Disposons maintenant de $A_1$ de la manière suivante. Soit $t_0$ une valeur positive aussi grande qu'on veut et $A_1 < \frac{1}{2}$ tel que pour $A > A_1$

$$\lambda |\eta(A)| < \frac{1}{t_0} < 1, \quad |\rho(A)| < \frac{1}{2t_0^2} < \frac{1}{2}, \quad B(A) > \frac{1}{2}. \tag{34}$$

Un tel choix de $A_1$ est toujours possible à raison des propriétés supposées des fonctions $\eta(A)$, $\rho(A)$ et $B(A)$. Les $A_1$ et $t_0$ donnés, on cherche un $n_1$ tel que pour $n > n_1$

$$t_1 = nr(1 - A_1) > t_0, \quad \frac{(\lambda+1)\,t_0}{nr} < \frac{1}{t_0}. \tag{35}$$

Utilisant le fait que, pour une fonction dérivable $F(x)$ et des valeurs assez petites de $|h|$, on a toujours

$$|F(x+h) - F(x)| < 2\,|h F'(x)| \tag{36}$$

on déduit de (29) et (28) (en regardant $\varphi$ comme fonction de $A$, pour donné)

$$|\varphi(t)| < 2e^{-t} t^{m-1} \left\{ |\rho t| + y|\eta| + x\left|\frac{t}{nr}\right| + y\left|\frac{t\lambda(1+\eta)}{nr}\right| \right\}, \tag{37}$$

donc pour $t < t_0$ suivant (34) et (35):

$$|\varphi(t)| < 2e^{-t} t^{m-1} \frac{1}{t_0} \left\{ \frac{1}{2} + \frac{1}{\lambda} + x + y \right\} < \frac{c}{t_0} e^{-t} t^{m-1}, \tag{38}$$

et en intégrant:

$$\int_0^{t_0} |\varphi|\, dt < \frac{c}{t_0} \int_0^{\infty} e^{-t} t^{m-1} dt = \frac{c\,(m-1)!}{t_0}. \tag{39}$$

Donc, on peut diminuer le deuxième terme de (32) autant qu'on veut, en augmentant $t_0$.

Pour le troisième terme constatons, qu'à raison de $|\lambda\eta| < 1$, $|\rho| < \frac{1}{2}$, $A > \frac{1}{2}$, $B > \frac{1}{2}$ dans tout l'intervalle de 0 à $t_1$ il suit de (28):

$$\int_{t_0}^{t_1} f\, dt < \int_{t_0}^{t_1} e^{-\frac{t}{2}} t^{m-1} \left(1 + \frac{1}{\lambda}\right)^y \left(\frac{1}{2}\right)^{-x} \left(\frac{1}{2}\right)^{-y} dt < 2^m \left(1 + \frac{1}{\lambda}\right)^y \int_{t_0}^{\infty} e^{-\frac{t}{2}} t^{m-1}\, dt. \tag{40}$$

Comme l'intégrale au dernier membre est convergente, sa valeur diminue vers zéro, si $t_0$ augmente. Il est de même pour le quatrième terme de (32). Alors l'équation (25) est établie.

La deuxième intégrale dans (6) ne diffère de la première que par l'échange de $A$ en $B$, de $x$ en $y$ et de $\alpha$ en $\beta$. Il faut donc remplacer $\lambda$ par $\frac{1}{\lambda}$ et $r = \alpha + \beta\lambda$ par $\beta + \lambda^{-1}\alpha = r\lambda^{-1}$. Donc, il suit de (25)

$$\lim_{n\to\infty} n^m I_B = (m-1)!\, r^{-m} \lambda^m \cdot \lambda^{-x} = (m-1)!\, r^{-m} \lambda^{-y} = \lim_{n\to\infty} n^m I_A. \tag{41}$$

Or, on a d'après (6):

$$\lim_{n\to\infty} P(x) = \lim_{n\to\infty} \binom{\alpha n}{x} \binom{\beta n}{y} (x I_A + y I_B) = \frac{\alpha^x \beta^y}{x!\,y!} \lim_{n\to\infty} n^m (x I_A + y I_B),$$

et enfin en utilisant (25) et (41), puisque $x+y=m$

$$\lim_{n\to\infty} P(x)=\frac{\alpha^x\beta^y}{x!\,y!}\,m!\,r^{-m}\lambda^y=\binom{m}{x}\frac{\alpha^x(\lambda\beta)^y}{(\alpha+\lambda\beta)^m}\,. \tag{42}$$

Il est facile d'interpréter ce résultat, en le comparant avec (17), le résultat spécial pour le cas $A=B$. Les grandeurs $\alpha$, $\beta$ sont maintenant remplacées par $\alpha:(\alpha+\beta\lambda)$ et $\beta\lambda:(\alpha+\beta\lambda)$. Le nombre $\lambda$, défini par (21), peut être considéré comme le rapport des probabilités intitiales „des plus grandes valeurs $z$" dans les deux classes $A$ et $B$. La probabilité cherchée $P(x)$ correspond à celle d'épreuves $m$ fois répétées dans l'alternative de deux événements, dont les probabilités élémentaires sont en rapport $\alpha$ à $\beta\lambda$. La seule modification apportée au résultat spécial (17) consiste dans l'échange du rapport $\alpha:\beta$ (rapport des étendues des deux classes) par le rapport

$$\alpha:\lambda\beta=\lim_{A\to 1,\ B\to 1}[\alpha(1-A):\beta(1-B)] \tag{43}$$

dans lequel interviennent les probabilités initiales „des plus grandes valeurs $z$" dans les deux classes.

### 4. Cas général $A\neq B$. Limite pour $n$ infini, $\mu=m:n$ fini

Soit maintenant

$$m=\mu n,\quad x=\xi n,\quad y=\eta n,\quad \mu=\xi+\eta. \tag{44}$$

Nous cherchons la limite de $P(x)$ pour $n$ infini quand $\mu$ reste constant, $\xi$ et $\eta$ restant bornés. Désignons par $P_A(x)$ la première partie de l'expression (6) pour $P(x)$, c'est-à-dire:

$$P_A(x)=\binom{\alpha n}{\xi n}\binom{\beta n}{\eta n}\xi n\int_0^1[A^{\alpha-\xi}(1-A)^\xi B^{\beta-\eta}(1-B)^\eta]^n\frac{dA}{1-A}\,. \tag{45}$$

D'après la formule bien connue de Stirling il suit de (45):

$$\lim_{n\to\infty}\sqrt{n}\,P_A(x)=\frac{1}{2\pi}\lim_{n\to\infty}\frac{\xi\sqrt{\alpha\beta}}{\sqrt{\xi(\alpha-\xi)\eta(\beta-\eta)}}\sqrt{n}\int_0^1 f^n(A,\xi)\frac{dA}{1-A}\,, \tag{46}$$

où

$$f(A,\xi)=\left(\frac{\alpha A}{\alpha-\xi}\right)^{\alpha-\xi}\left(\frac{\alpha(1-A)}{\xi}\right)^\xi\left(\frac{\beta B}{\beta-\eta}\right)^{\beta-\eta}\left(\frac{\beta(1-B)}{\eta}\right)^\eta. \tag{47}$$

Ici le deuxième membre de (46) définit une fonction continue de $\xi$, dont les valeurs pour les valeurs entières de $n\xi$ coïncident avec la limite de

$$\sqrt{n}\,P_A(n\xi).$$

Remarquons tout d'abord qu'il existe toujours une et une seule valeur de $A$ respectivement $B$, pour laquelle chacun des quatre quotients dans (47) prend la valeur 1. En effet l'équation

$$\mu=1-\alpha A(z)-\beta B(z),\quad 0<\mu<1, \tag{48}$$

où le second membre est une fonction décroissante de 1 à 0, possède au moins une solution $z_0$. Il y en a plus qu'une (une infinité de solutions) si

$\alpha A+\beta B$ reste constant à l'environ de $z_0$. Mais, comme $A(z)$, $B(z)$ sont monotones, $\alpha$ et $\beta$ — positifs, $\alpha A+\beta B$ ne peut rester constant que si $A(z)$ et $B(z)$ sont constants tous les deux. Donc, il n'existe qu'une seule valeur $A_0$ et une seule valeur $B_0$ satisfaisant à l'équation (48).

Les valeurs

$$\begin{aligned}\xi_0 &= \alpha(1-A_0),\\ \eta_0 &= \beta(1-B_0),\\ \xi_0+\eta_0 &= 1-\alpha A_0-\beta B_0=\mu\end{aligned} \tag{49}$$

rendent chacun des quatre quotients au deuxième membre de (47) égal à l'unité. Nous supposerons dans ce qui suit que la fonction $B(A)$ possède une dérivée $B_0'$ finie et différente de zéro au point $A=A_0$.

Or, pour évaluer la limite (46) nous nous appuyerons sur le théorème suivant, légère généralisation d'un théorème que j'ai établi en 1919 [2]. Soit $f(A, \xi)$ une fonction intégrable (relative à $A$) dans $A_1<A<A_2$ et $\xi_1<\xi<\xi_2$, jouissant des propriétés suivantes:

a) pour $A=A_0$, $\xi=\xi_0$ il est

$$\left.\begin{aligned} f=1,\ \frac{\partial f}{\partial A}=\frac{\partial f}{\partial \xi}=0,\ \frac{\partial^2 f}{\partial A^2}=-L<0,\ \frac{\partial^2 f}{\partial\xi\partial A}=-M,\ \frac{\partial^2 f}{\partial \xi^2}=-N,\\ \frac{LN-M^2}{L}=h^2>0;\end{aligned}\right\} \tag{50}$$

b) les dérivées du troisième ordre sont bornées au voisinage du point $A=A_0$, $\xi=\xi_0$;

c) partout, sauf le point $A_0$, $\xi_0$, on a $|f|<1$. Soit de plus $\psi(\xi, A)$ une fonction continue au point $A_0$, $\xi_0$ et telle que $f^k\psi$ est borné pour un entier $k>0$; il suit alors, si $A_1<A_0<A_2$,

$$\lim_{n\to\infty} \sqrt{n}\int_{A_1}^{A_2} f^n\left(A,\ \xi_0+\frac{u}{\sqrt{n}}\right)\psi\left(A,\ \xi_0+\frac{u}{\sqrt{n}}\right)dA=\psi(A_0,\ \xi_0)\sqrt{\frac{2\pi}{L}}\,e^{-\frac{h^2u^2}{2}} \tag{51}$$

où la convergence est uniforme dans chaque intervalle fini de $u$.

Pour les détails de la démonstration je renvoie à mes publications précédentes et je ne donne ici que les traits principaux du calcul. Le théorème de Taylor nous fournit pour le logarithme naturel de $f$ la formule:

$$\begin{aligned}\ln f=-\frac{1}{2}\,[L(A-A_0)^2+2M(A-A_0)(\xi-\xi_0)+N(\xi-\xi_0)^2]+\\ +\text{termes de 3-ème degrès.}\end{aligned} \tag{52}$$

Les conditions a) y sont déjà prises en considération, la condition b) nous permet, en introduisant $t$ et $u$ par les formules

$$A-A_0=\frac{t}{\sqrt{n}},\quad \xi-\xi_0=\frac{u}{\sqrt{n}}, \tag{53}$$

d'en tirer la conséquence

$$\lim_{n\to\infty}\ln f^n\left(A_0+\frac{t}{\sqrt{n}},\ \xi_0+\frac{u}{\sqrt{n}}\right)=-\frac{1}{2}(Lt^2+2Mtu+Nu^2). \tag{54}$$

[2] Voir mon mémoire „Fundamentalsätze der Wahrscheinlichkeitsrechnung", inséré au „Mathematische Zeitschrift", t. 4, (1919), p. 1—96, surtout § 1 et 2; plus tard: „Vorlesungen aus dem Gebiete der angewandten Mathematik", T. 1. „Wahrscheinlichkeitsrechnung", Leipzig und Wien, 1931, § 8.

La convergence est uniforme dans chaque intervalle fini de $t$ et $u$. Mais la condition c) et les propriétés supposées de $\psi$ rendent possible d'étendre l'endroit où (54) subsiste, de sorte que l'intégration de $u=-\infty$ à $u=\infty$ est permise. Les détails de cette déduction sont analogues à ceux du paragraphe précédent. On a donc

$$\lim_{n\to\infty} \sqrt{n} \int_{A_1}^{A_2} f^n \psi\, dA = \psi(A_0, \xi_0) \int_{-\infty}^{\infty} e^{-\frac{1}{2}(Lt^2+2Mtu+Nu^2)} dt =$$

$$= \sqrt{\frac{2\pi}{L}}\, \psi(A_0, \xi_0)\, e^{-\frac{1}{2}h^2u^2}. \qquad (55)$$

Dans notre cas spécial la fonction $f(A, \xi)$ est définie par (47). Les dérivées du premier ordre sont:

$$\left.\begin{aligned} \frac{\partial f}{\partial A} &= \left[\frac{\alpha-\xi}{A} - \frac{\xi}{1-A} + B'\left(\frac{\beta-\eta}{B} - \frac{\eta}{\lambda-B}\right)\right] f, \\ \frac{\partial f}{\partial \xi} &= \left[\ln \frac{1-A}{A}\frac{B}{1-B} + \ln \frac{\alpha-\xi}{\xi}\frac{\eta}{\beta-\eta}\right] f. \end{aligned}\right\} \qquad (56)$$

Les deux expressions s'annulent si $\xi$ et $\eta$ prennent les valeurs $\xi_0$, $\eta_0$ et $A$, $B$ les valeurs $A_0$, $B_0$, indiquées en (49). Les trois dérivées du deuxième ordre au point $A=A_0$, $\xi=\xi_0$ se trouvent:

$$L = \frac{\alpha^3}{\xi_0(\alpha-\xi_0)} + B_0'^2 \frac{\beta^3}{\eta_0(\beta-\eta_0)}, \quad M = \frac{\alpha^2}{\xi_0(\alpha-\xi_0)} - B_0' \frac{\beta^2}{\eta_0(\beta-\eta_0)},$$

$$N = \frac{\alpha}{\xi_0(\alpha-\xi_0)} + \frac{\beta}{\eta_0(\beta-\eta_0)}, \qquad (57)$$

d'où on déduit facilement:

$$h^2 = \frac{LN-M^2}{L} = \frac{\alpha\beta(\alpha+\beta B_0')^2}{\alpha^3\eta_0(\beta-\eta_0) + B_0'^2\beta^3\xi_0(\alpha-\xi_0)}. \qquad (58)$$

Les conditions a) et b) sont donc remplies; pour c) remarquons que le produit des deux premiers facteurs de (47) reste au-dessous de l'unité si $\xi \neq \alpha(1-A)$ et le produit de deux autres si $\eta \neq \beta(1-B)$.

La fonction $\psi(A, \xi)$ est donnée par

$$\psi(A, \xi) = \frac{1}{2\pi} \frac{\xi\sqrt{\alpha\beta}}{\sqrt{\xi(\alpha-\xi)\eta(\beta-\eta)}} \frac{1}{1-A}. \qquad (59)$$

On voit bien qu'elle remplit les conditions du théorème, puisque $f^k\psi$ est borné pour $k>\xi^{-1}$, la fonction $f$ comprenant le facteur $(1-A)^\xi$. Nous avons donc, en remplaçant $\xi_0:(1-A_0)$ par $\alpha$

$$\lim_{n\to\infty} \sqrt{n}\, P_A(n\xi_0 + \sqrt{n}u) = \frac{\alpha\sqrt{\alpha\beta}}{\sqrt{2\pi L}} \frac{1}{\sqrt{\xi_0(\alpha-\xi_0)\eta_0(\beta-\eta_0)}} e^{-\frac{1}{2}h^2u^2}. \qquad (60)$$

La deuxième partie $P_B$ de l'expression (6) ne diffère de la première que par l'échange de $A$ et $B$ etc. On voit bien que ça revient à remplacer $dA$ dans l'intégrale (55) par $B_0'dA$ et le premier facteur $\alpha$ dans (60) par $\beta$. Il faut donc finalement, pour l'expression $P(x)$, remplacer le coefficient $\alpha$ de (60) par

$\alpha+\beta B_0'$. La comparaison avec (57) et (58) nous fournit le résultat définitif:

$$\lim_{n\to\infty} \sqrt{n}\, P(n\xi_0+\sqrt{n}u) = \frac{h}{\sqrt{2\pi}} e^{-\frac{1}{2}h^2u^2}, \tag{61}$$

la „précision“ $h$ étant donnée par (58). La convergence est uniforme dans tout l'intervalle infini de $u$. Le résultat est à préciser en ce sens, que $P(x)$ dans (61) signifie une fonction d'une variable indépendante continue, fonction qui, pour des valeurs entières de $x$, coïncide avec la probabilité cherchée $P(x)$. Pour l'application pratique on déduit de (61) comme formule approchée, valable pour des grandes valeurs de $n$:

$$P(x) \approx \frac{h}{\sqrt{2\pi n}} e^{-\frac{1}{2}\frac{h^2(x-n\xi_0)^2}{n}}. \tag{62}$$

Au cas $A=B$, où $B_0'=1$, la dernière équation (49) donne $\mu=1-A_0$ d'où on déduit $\xi_0=\alpha\mu$, $\eta_0=\beta\mu$ de sorte qu'on revient au résultat (20).

## 5. Le problème inverse

Pour pouvoir faire l'analyse d'un ensemble d'observations données, concernant un cas spécial d'une statistique de deux races, il faut encore examiner le problème „inverse“. Il s'agit, d'une façon générale, de la question: étant données les valeurs de $x$ et $y$, quelle est la probabilité pour que ces valeurs résultent d'une certaine relation entre les distributions $A$ et $B$? Cette question est analogue au problème bien connu dit de Beyes, problème inverse à celui de Bernoulli. Nous y allons jeter un coup d'œil rapide.

Supposons qu'il s'agit du premier cas asymptotique, où $n$ tend vers l'infini, tandis que $m$, $x$, $y$ restent finis. Nous savons déjà qu'en ce cas les distributions $A(z)$ et $B(z)$ n'entrent dans le résultat que par le rapport

$$\lambda = \lim_{A\to 1} \frac{B-1}{A-1}$$

qui donc est la seule valeur exerçant une influence sur $P(x)$. Pour fixer les idées admettons maintenant que la relation entre $A$ et $B$ soit d'une forme simple donnée par exemple par les équations

$$\left.\begin{array}{lll} 1-B=\lambda(1-A) & \text{pour} & A+B\geqq 1, \\ B=\frac{1}{\lambda}A & \text{„} & A+B\leqq 1. \end{array}\right\} \tag{63}$$

Ces équations définissent une fonction continue croissante $B(A)$, représentée par deux lignes droites, se rencontrant sur la droite $A+B=1$ et symétriques à cette dernière. L'expression donnée en (6) pour $P(x)$ est rendue par (63), fonction univoque de $\lambda$, fonction que nous allons signifier par $Q(\lambda)$. Soit maintenant $q(\lambda)\,d\lambda$ la probabilité „a priori“ pour que la valeur de $\lambda$ tombe dans un certain intervalle $\lambda$, $\lambda+d\lambda$; on aura pour la probabilité „a posteriori“:

$$p(\lambda) = \text{const}\cdot q(\lambda)\, Q(\lambda), \tag{64}$$

où la constante se détermine moyennant la condition

$$\int_0^\infty p(\lambda)\, d\lambda = 1. \tag{65}$$

Une difficulté s'élève à l'égard de $q(\lambda)$. Évidemment il n'y a pas de raison d'admettre $q(\lambda) = \text{const}$, puisqu'on pourrait remplacer $\lambda$ par $1:\lambda$ et, après ce changement, demander à juste titre que la distribution des probabilités sur l'axe des valeurs $1:\lambda$ soit encore uniforme, — ce qui serait incompatible avec $q(\lambda) = \text{const}$. Une hypothèse plus raisonnable se présente si l'on regarde l'angle même (au lieu de sa tangente) formé par la droite $1 - B = \lambda(1 - A)$ et l'axe des abscisses comme la variable à distribution uniforme des probabilités a priori, ce qui revient à $q(\lambda) = \text{const} : (1 + \lambda^2)$. On pourrait, en adoptant cette hypothèse, évaluer $p(\lambda)$ à l'aide de (63), toujours avec $Q(\lambda) = P(x)$, soit qu'on se sert de l'expression complète (6) ou de l'approximation asymptotique (42) pour $P(x)$. Mais, en tout cas, on ne peut pas attacher trop d'importance à une solution d'un problème inverse autant que le résultat dépend d'une façon décisive du choix, arbitraire à une certaine mesure, de la probabilité a priori.

La théorie du problème de Beyes nous démontre que l'influence de la probabilité a priori s'évanouit peu à peu si le nombre des cas envisagés augmente [3]. Il en suit que, dans la question actuelle, pour une valeur trop petite de $m$ les conclusions concernant la probabilité a posteriori de $\lambda$ ne seront pas assez exactes, tandis qu'on arrivera, pour $n$ assez grand, à un résultat légitime, en supposant n'importe quelle loi des probabilités a priori. Substituons donc pour $\lambda$ une nouvelle variable $z$, définie par

$$z = \frac{\alpha}{\alpha + \lambda\beta}, \qquad 1 - z = \frac{\lambda\beta}{\alpha + \lambda\beta} \tag{66}$$

et admettons que $q(\lambda)$ soit donnée par

$$q(\lambda)\, d\lambda = \text{const} \cdot dz. \tag{67}$$

De cette façon on déduit de (64) moyennant (42)

$$p(\lambda)\, d\lambda = \text{const} \cdot z^x (1 - z)^{m-x}\, dz \tag{68}$$

et d'après (65), utilisant une formule d'intégration bien connue

$$p(\lambda)\, d\lambda = (m + 1) \binom{m}{x} z^x (1 - z)^{m-x}\, dz. \tag{69}$$

Cette formule nous permet de déterminer la probabilité pour que $\lambda$ tombe dans des certaines limites $\lambda_1$, $\lambda_2$.

La valeur moyenne et la déviation de la distribution (69), si l'on regarde $z$ comme la variable indépendante, sont respectivement:

$$\frac{x + 1}{m + 1} \text{ et } \frac{(x + 1)(m - x + 1)}{(m + 2)^2 (m + 3)}. \tag{70}$$

L'inégalité bien connue de Tchebycheff nous fournit donc le résultat:

Il y a une probabilité supérieure à

$$1 - \frac{(x + 1)(m - x + 1)}{(m + 2)^2 (m + 3)} \frac{1}{\zeta^2} \tag{71}$$

---

[3] L'indépendance du résultat, dans le problème de Beyes, de la probabilité „a priori" pour un nombre assez grand d'épreuves était établie dans mon mémoire cité de 1919, § 9, et plus tard dans mon „Cours" de 1931, § 6. Voir aussi V. Romanovsky, „Giornale dell'Instituto Italiano degli Attuari", Anno 11, n. 4, ottobre 1931.

pour que $z$ tombe dans les limites

$$\frac{x+1}{m+1} - \zeta \quad \text{et} \quad \frac{x+1}{m+1} + \zeta. \tag{72}$$

Ce résultat, bien qu'il est basé sur l'hypothèse (67), en devient de plus en plus indépendant, si $m$ augmente.

N'insistant pas sur une étude, d'ailleurs analogue, du deuxième cas asymptotique qui était traité en n⁰ 4, nous passons à quelques applications.

## 6. Applications

Prenons d'abord le cas, où les deux distributions données sont des gaussiennes:

$$A(z) = \Phi(z), \quad B(z) = \Phi[k(z-a)],$$

où $\Phi$ signifie, comme d'habitude,

$$\Phi(z) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{z} e^{-x^2}\, dx.$$

Ici les conditions supposées sont remplies et, par élimination de $z$, on trouve $B$ comme fonction continue, monotone croissante et différentiable de $A$. Pour $A$ tendant vers l'unité on a

$$\lambda = \lim \frac{B-1}{A-1} = \lim \frac{dB}{dA} = \lim_{z \to \infty} k e^{-k^2(z-a)^2 - z^2}.$$

On voit donc que

$$\begin{aligned} \lambda &= \infty \text{ pour } k > 1 \text{ ou } k = 1 \text{ et } a > 0, \\ \lambda &= 0 \quad ,, \quad k < 1 \quad ,, \quad k = 1 \quad ,, \quad a < 0. \end{aligned}$$

Or, il suit des résultats du n⁰ 3:

Si le caractère distinctif dans chacune de deux classes, différente l'une de l'autre, obéit à une loi de Gauss, et si, pour $n$ très grand, on ne considère qu'un nombre restreint des plus grandes valeurs $z$, il sera presque sûr que seulement les individus d'une seule des deux classes figurent parmi les valeurs considérées.

Passons à un autre exemple. Dans un pays d'Europe, dont le nombre d'habitants est d'environ $65 \cdot 10^6$, la population se compose de deux races aux chiffres relatifs $\alpha = 0{,}009$, $\beta = 0{,}991$. Un tout petit nombre de ces habitants s'occupent des recherches scientifiques dans les domaines de la physique et de la chimie. Une échelle absolue pour mesurer la capacité scientifique (la valeur $z$) d'un physicien ou d'un chimiste n'existe pas. Toutefois on pourra admettre que les titulaires du prix Nobel constituent un ensemble des plus grandes valeurs du caractère distinctif $z$. La liste des titulaires des années 1901 à 1933 comprend 27 noms originaires du dit pays, parmi lesquels 5 appartenant à la classe $A$, de sorte qu'on a $m = 27$, $x = 5$, $y = 22$. Confrontons d'abord ces chiffres avec la formule (18) en admettant $n = \infty$ et supposant que les deux distributions des probabilités $A$ et $B$ soient égales entre elles. On trouve

$$P(5) = \binom{27}{5} 0{,}009^5 \cdot 0{,}991^{22} = 0{,}00000391.$$

Comme il y a 28 valeurs possibles de $x$, cette probabilité est à peu près dix mille fois inférieure à la probabilité moyenne. La probabilité pour que le nombre $x$ soit supérieur à zéro est

$$1-P(27)=1-0{,}991^{27}=0{,}2166.$$

La probabilité pour que $x$ soit supérieur à l'unité est au-dessous de 0,03, d'ailleurs la probabilité d'une valeur de $x$ inférieure à 5 ne diffère de l'unité que de $10^{-5}$ près. Il suit de tout ça que — sous la condition de $A$ et $B$ égales — le résultat $x=5$ qui s'est produit dans l'expérience, serait d'une probabilité extrêmement petite.

On s'en peut douter si „le plus grand physicien" se trouve ou non parmi les noms appartenant à la première des deux classes. Annotons, en tout cas, que la probabilité du premier fait — l'égalité $A=B$ supposée — serait d'après (17) ou (7), avec $m=1$, $x=1$:

$$P_1(1)=\alpha=0{,}009.$$

D'autre part, en rejetant l'hypothèse $A=B$, nous parviendrons à des probabilités bien plus raisonnables pour le fait observé $x=5$, $y=22$. Admettons d'abord que $\lambda$ satisfasse à l'équation

$$\alpha:\lambda\beta=x:y=5:22,\quad \lambda=0{,}04=\frac{1}{25}.$$

C'est donc l'hypothèse que la probabilité d'un très grand talent en physique ou chimie est appréciée 25 fois plus grande pour un individu de la classe $A$ que pour un tel de la classe $B$. Sous cette condition on a d'après (42) avec $\alpha=0{,}009$, $\lambda\beta=0{,}0396$:

$$P(5)=\binom{27}{5}0{,}009^5\cdot0{,}0396^{22}\cdot0{,}0486^{-27}=0{,}194.$$

Cette valeur est environ 5,5 fois plus grande que la probabilité moyenne 1 : 28. Le résultat $x=5$ constitue, maintenant, le cas le plus probable. La probabilité pour que $x$ reste dans les limites 3 à 7 se trouve, avec $\lambda=0{,}04$, égale à 0,79. Il est donc évident que l'hypothèse $\lambda=0{,}04$ mène à des chiffres conformes aux observations.

Enfin, on peut se servir des formules établies dans n° 5 pour le problème inverse. Le nombre $m=27$ est assez grand pour en justifier l'usage. La moyenne et la déviation de la distribution (69) se déterminent d'après (70) comme il suit:

$$\frac{x+1}{m+1}=\frac{6}{28}=0{,}214,\quad \frac{(x+1)(m-x+1)}{(m+2)^2(m+3)}=\frac{6\cdot23}{29^2\cdot30}=0{,}000547.$$

La transformation (66) s'écrit:

$$z=\frac{\alpha}{\alpha+\lambda\beta},\quad \lambda=\frac{\alpha}{\beta}\frac{1-z}{z}=\frac{9}{991}\frac{1-z}{z}.$$

Donc, à la valeur moyenne $z_0=0{,}214$ correspond la valeur $\lambda_0=0{,}33$. Prenons maintenant $\xi=0{,}06$, nous aurons

$$z_1=z_0-\xi=0{,}154,\quad z_2=z_0+\xi=0{,}274$$

et les valeurs correspondantes de $1:\lambda$ seront:

$$\frac{1}{\lambda_1}=20,1, \quad \frac{1}{\lambda_2}=41,6,$$

tandis que (71) nous donne

$$1-\frac{0,000547}{0,06^2}=1-0,152=0,848.$$

Alors il y a une probabilité d'à peu près 85% pour le fait que, parmi les individus de la classe $A$, la probabilité d'un talent éminent en physique ou chimie soit au moins 20 fois et au plus 42 fois plus grande que dans la classe $B$.

(Поступило в редакцию 20/V 1934 г.)

---

# Проблема двух рас

**Р. Мизес** (Стамбул)

Нижеследующая задача часто встречается в математической статистике и ее приложениях к биологии и страховому делу.

Дана совокупность из $n$ индивидов, распределяющихся на два класса $A$ и $B$ так, что $\alpha n$ из них принадлежат классу $A$, $\beta n$ — классу $B$. Пусть $z$ — некоторый количественный признак элемента совокупности, например его длина, и пусть для каждого из двух классов существует функция, выражающая распределение вероятностей, $A(z)$ [соответственно, $B(z)$]. В частности, $A(z)$ выражает вероятность того, что количественная характеристика элемента класса $A$ не превзойдет $z$.

Требуется узнать вероятность $P(x)$ того, что среди $m$ индивидов, отвечающих наибольшим значениям $z$, найдутся $x$ индивидов класса $A$ и $y=m-x$ класса $B$.

Очевидно, что и в том случае, когда отсутствуют способы, позволяющие измерить значения $z$, вопрос имеет смысл. Достаточно для его постановки, если мы умеем упорядочить элементы, расположив их в порядке возрастания $z$, так, что случай $z_1>z_2$ совершенно отличается от случая $z_1<z_2$.

Мы можем применить полученные результаты и тогда, когда статистика задает нам такое разбиение совокупности на две группы, при котором величины $z$ одной из них больше значений $z$ другой. Так, например, мы склонны допустить, что обладатели нобелевской премии среди физиков какой-нибудь страны относятся к „самым крупным" представителям этой дисциплины. Тогда, если обитатели этой страны принадлежат к двум различающимся между собой расам, — наши изыскания будут применимы.

На следующих страницах мы даем сначала ($n^0$ 1) общее выражение для $P(x)$ в виде некоторого стильтьесова интеграла (6), сводящегося к обычному определенному интегралу при непрерывности $A(z)$ и $B(z)$. За-