

Une loi-limite du calcul des probabilités.

Par E. FELDHEIM à Budapest.

1. Désignons par p la probabilité constante d'un évènement fortuit E , et soit m la répétition de l'évènement E au cours de n épreuves (c'est-à-dire le nombre de ces épreuves où E s'est produit). On fait $q = 1 - p$.

La loi-limite de LAPLACE dit que la probabilité de l'inégalité

$$t_1 < \frac{m - np}{\sqrt{npq}} < t_2$$

tend, pour $n \rightarrow \infty$, vers

$$(1) \quad P(t_1, t_2) = \frac{1}{\sqrt{2\pi}} \int_{t_1}^{t_2} e^{-\frac{z^2}{2}} dz,$$

où t_1 et $t_2 \geq t_1$ sont deux nombres réels arbitraires.

M. KHINTCHINE a fait l'étude du cas où t_1 et t_2 deviennent infiniment grands en même temps, et il a trouvé le théorème suivant:¹⁾

Si l'on pose $t_1 = t + \frac{g_1}{t}$, $t_2 = t + \frac{g_2}{t}$, où $0 \leq g_1 \leq g_2 \leq +\infty$, on aura à la limite

$$(2) \quad \frac{P(t_1, t_2)}{P(t, \infty)} = e^{-g_1} - e^{-g_2},$$

pour $t \rightarrow \infty$, $t = o(\sqrt{n})$.

Nous donnerons d'abord une démonstration élémentaire mais non rigoureuse de ce théorème.

¹⁾ A. KHINTCHINE, Über einen neuen Grenzwertsatz der Wahrscheinlichkeitsrechnung, *Math Annalen*, 101 (1929), p. 745—752.

Si $t \rightarrow \infty$, le rapport

$$\frac{P(t_1, t_2)}{P(t, \infty)} = \frac{\int_{t_1}^{t_2} e^{-\frac{z^2}{2}} dz}{\int_t^{\infty} e^{-\frac{z^2}{2}} dz} = \frac{\int_{t+\frac{g_2}{t}}^{t+\frac{g_1}{t}} e^{-\frac{z^2}{2}} dz}{\int_t^{\infty} e^{-\frac{z^2}{2}} dz}$$

n'a aucun sens. On peut appliquer la règle de L'HÔPITAL :

$$\begin{aligned} \frac{P(t_1, t_2)}{P(t, \infty)} \Big|_{t=\infty} &= \frac{\left(1 - \frac{g_2}{t^2}\right) e^{-\frac{1}{2}\left(t+\frac{g_2}{t}\right)^2} - \left(1 - \frac{g_1}{t^2}\right) e^{-\frac{1}{2}\left(t+\frac{g_1}{t}\right)^2}}{-e^{-\frac{t^2}{2}}} \Big|_{t=\infty} \\ &= \left[\left(1 - \frac{g_1}{t^2}\right) e^{-g_1} \cdot e^{-\frac{g_1^2}{2t^2}} - \left(1 - \frac{g_2}{t^2}\right) e^{-g_2} \cdot e^{-\frac{g_2^2}{2t^2}} \right]_{t=\infty} \\ &= \left[e^{-g_1} \left[1 + O\left(\frac{g_1}{t^2}\right) \right] - e^{-g_2} \left[1 + O\left(\frac{g_2}{t^2}\right) \right] \right]_{t=\infty} = e^{-g_1} - e^{-g_2}. \end{aligned}$$

2. Nous donnerons une généralisation de ce théorème au cas des „tirages contagieux“ de M. PÓLYA.

Avant de formuler notre théorème, considérons le cas suivant, analogue à celui de M. KHINTCHINE. Soit $P(t)$ la probabilité pour qu'un certain écart absolu fortuit X surpasse un nombre positif t . On suppose que

$$(3) \quad P(t) = A \int_t^{\infty} e^{-\lambda x^r} dx,$$

où A, λ, r sont certaines constantes positives.

Pour déterminer ces constantes, observons que $P(0) = 1$, c'est-à-dire

$$A \int_0^{\infty} e^{-\lambda x^r} dx = 1$$

d'où l'on tire

$$A = \frac{r \lambda^{\frac{1}{r}}}{\Gamma\left(\frac{1}{r}\right)},$$

$\Gamma(x)$ étant l'intégrale eulérienne de seconde espèce.

Pour calculer λ , appelons μ le moment d'ordre r de l'écart x ,

$$\mu = A \int_0^{\infty} x^r e^{-\lambda x} dx = \frac{1}{\lambda r},$$

donc

$$(4) \quad A = \frac{r^{1-\frac{1}{r}}}{\mu^{\frac{1}{r}} \Gamma\left(\frac{1}{r}\right)}, \quad \lambda = \frac{1}{\mu r}.$$

La probabilité de l'inégalité $t_1 \leq x \leq t_2$ sera alors

$$(4') \quad P(t_1, t_2) = A \int_{t_1}^{t_2} e^{-\frac{1}{\mu r} x^r} dx.$$

Examinons le cas où t_1 et t_2 tendent vers l'infini tous les deux, en posant

$$t_1 = t + \frac{a}{t^{r-1}}, \quad t_2 = t + \frac{b}{t^{r-1}} \quad (r > 1, 0 < a < b),$$

$$\begin{aligned} P(t_1, t_2) &= A \int_{t+\frac{a}{t^{r-1}}}^{t+\frac{b}{t^{r-1}}} e^{-\frac{1}{\mu r} x^r} dx = \frac{A}{t^{r-1}} \int_a^b e^{-\frac{1}{\mu r} \left(t + \frac{y}{t^{r-1}}\right)^r} dy = \\ &= \frac{A e^{-\frac{t^r}{\mu r}}}{t^{r-1}} \int_a^b e^{-\frac{y}{\mu}} dy \left\{ 1 + O\left(\frac{1}{t^r}\right) \right\} = \{1 + O(1)\} \left(e^{-\frac{a}{\mu}} - e^{-\frac{b}{\mu}} \right) \frac{A \mu e^{-\frac{t^r}{\mu r}}}{t^{r-1}} = \\ &= \{1 + O(1)\} \left(e^{-\frac{a}{\mu}} - e^{-\frac{b}{\mu}} \right) \frac{(\mu r)^{1-\frac{1}{r}} e^{-\frac{t^r}{\mu r}}}{t^{r-1} \Gamma\left(\frac{1}{r}\right)}. \end{aligned}$$

De même,

$$\begin{aligned} P(t) &= A \int_t^{\infty} e^{-\frac{1}{\mu r} x^r} dx = \frac{A}{t^{r-1}} \int_0^{\infty} e^{-\frac{1}{\mu r} \left(t + \frac{y}{t^{r-1}}\right)^r} dy, \\ P(t) &= \frac{A e^{-\frac{t^r}{\mu r}}}{t^{r-1}} \int_0^{\infty} e^{-\frac{y}{\mu}} dy \left\{ 1 + O\left(\frac{1}{t^r}\right) \right\} = \{1 + O(1)\} \frac{(\mu r)^{1-\frac{1}{r}} e^{-\frac{t^r}{\mu r}}}{t^{r-1} \Gamma\left(\frac{1}{r}\right)}, \end{aligned}$$

donc

$$(5) \quad \lim_{t \rightarrow \infty} \frac{P(t_1, t_2)}{P(t)} = e^{-\frac{a}{\mu}} - e^{-\frac{b}{\mu}}.$$

3. Considérons une urne contenant des boules blanches et noires, sa composition avant tout tirage étant donnée par p et q . La méthode des „tirages contagieux“ consiste non seulement en remettant la boule tirée, mais en mettant encore dans l'urne un nombre déterminé de boules de la même espèce. Supposons par exemple, qu'on met un multiple γ du nombre total des boules. Ce nombre γ sera appelé „facteur de contagion“.

La probabilité de tirer m boules blanches dans n épreuves sera donnée par la formule

$$(6) \quad P_m = \frac{n!}{m!(n-m)!} \cdot \frac{p(p+\gamma) \dots [p+(m-1)\gamma] q(q+\gamma) \dots [q+(n-m-1)\gamma]}{1(1+\gamma) \dots [1+(n-1)\gamma]}.$$

On peut mettre cette expression sous la forme plus condensée :

$$(7) \quad P_m = \frac{n}{m(n-m)} \frac{\Gamma(n)}{\Gamma(m)\Gamma(n-m)} \cdot \frac{\Gamma\left(\frac{p}{\gamma} + m\right) \Gamma\left(\frac{q}{\gamma} + n - m\right) \Gamma\left(\frac{1}{\gamma}\right)}{\Gamma\left(\frac{1}{\gamma} + n\right) \Gamma\left(\frac{p}{\gamma}\right) \Gamma\left(\frac{q}{\gamma}\right)}.$$

Nous démontrons le théorème suivant :

Si m désigne la répétition des boules blanches dans les tirages contagieux de PÓLYA, le facteur de contagion étant γ , le quotient des probabilités des deux inégalités

$$\frac{n\gamma}{\mu r} \left(t + \frac{a}{t^{r-1}}\right)^r < m < \frac{n\gamma}{\mu r} \left(t + \frac{b}{t^{r-1}}\right)^r \quad (0 < a < b) \\ (r > 1)$$

et

$$\frac{n\gamma}{\mu r} t^r < m < \infty$$

tend vers

$$e^{-\frac{a}{\mu}} - e^{-\frac{b}{\mu}}$$

si $p = \frac{\gamma}{r}$, n et t tendant vers l'infini, et γ tendant vers 0. Le

nombre μ désigne la valeur probable de la variable

$$x = \frac{\mu r}{n\gamma} m.$$

a) Cherchons la limite vers laquelle tend la probabilité P_m , lorsque n augmentant indéfiniment, p et γ tendent vers 0 de sorte que $\frac{p}{\gamma}$ conserve une valeur finie $\frac{1}{r}$ et $n\gamma$ tend vers l'infini.

Posons alors

$$m = n\gamma\lambda x = \frac{n\gamma}{\mu r} x.$$

P_m prendra la forme suivante :

$$P_m = \frac{n}{\frac{n\gamma x}{\mu r} n \left(1 - \frac{\gamma x}{\mu r}\right)} \frac{\Gamma(n)}{\Gamma\left(n + \frac{1}{\gamma}\right)} \cdot \frac{\Gamma\left(n \frac{\gamma x}{\mu r} + \frac{1}{r}\right)}{\Gamma\left(n \frac{\gamma x}{\mu r}\right)} \frac{\Gamma\left[n \left(1 - \frac{\gamma x}{\mu r}\right) + \frac{1}{\gamma} - \frac{1}{r}\right]}{\Gamma\left[n \left(1 - \frac{\gamma x}{\mu r}\right)\right]} \frac{\Gamma\left(\frac{1}{\gamma}\right)}{\Gamma\left(\frac{1}{r}\right) \Gamma\left(\frac{1}{\gamma} - \frac{1}{r}\right)}.$$

En employant l'expression asymptotique de $\Gamma(u)$, on trouve, pour $n \rightarrow \infty$,

$$\frac{\Gamma(An + \alpha)}{\Gamma(An + \beta)} = \frac{(An + \alpha)^{An + \alpha - \frac{1}{2}} e^{-An - \alpha} \sqrt{2\pi} \{1 + O(1)\}}{(An + \beta)^{An + \beta - \frac{1}{2}} e^{-An - \beta} \sqrt{2\pi} \{1 + O(1)\}},$$

alors

$$(8) \quad \frac{\Gamma(An + \alpha)}{\Gamma(An + \beta)} = (An)^{\alpha - \beta} \{1 + O(1)\}.$$

On obtient alors pour P_m :

$$(9) \quad P_m = \frac{\mu r}{\gamma n} \frac{1}{(\mu r)^{\frac{1}{r}} \Gamma\left(\frac{1}{r}\right)} \frac{\gamma^{\frac{1}{r}} \Gamma\left(\frac{1}{\gamma}\right)}{\Gamma\left(\frac{1}{\gamma} - \frac{1}{r}\right)} x^{\frac{1}{r} - 1} \left(1 - \frac{\gamma x}{\mu r}\right)^{\frac{1}{\gamma} - \frac{1}{r} - 1} \{1 + O(1)\}.$$

Si $\gamma \rightarrow 0$, $\frac{1}{\gamma} \rightarrow \infty$,

$$\frac{\gamma^{\frac{1}{r}} \Gamma\left(\frac{1}{\gamma}\right)}{\Gamma\left(\frac{1}{\gamma} - \frac{1}{r}\right)} = 1 + O(\gamma),$$

$$\left(1 - \frac{\gamma x}{\mu r}\right)^{\frac{1}{\gamma} - \frac{1}{r} - 1} = e^{-r\gamma(x) + O(\gamma)},$$

avec

$$f_\gamma(x) = \frac{1}{\gamma} \sum_{k=0}^{\infty} \frac{1}{k+1} \left(\frac{\gamma x}{\mu r} \right)^{k+1}.$$

Donc

$$(10) \quad P_m = \frac{\mu r}{\gamma^n} \frac{1}{(\mu r)^{\frac{1}{r}} \Gamma\left(\frac{1}{r}\right)} x^{\frac{1}{r}-1} e^{-f_\gamma(x) + o(\gamma)}.$$

b) Donnons encore une autre méthode pour trouver la limite de la probabilité P_m , en faisant usage de la notion de fonction caractéristique. Cette méthode conduit à un résultat équivalent à celui du calcul direct, et, dans un grand nombre de cas, est plus rapide, et plus simple.

La fonction caractéristique est définie comme la valeur probable de la quantité e^{itx} , où t est un paramètre, donc, dans le cas des lois continues, cette fonction est

$$(11) \quad \varphi(t) = \int_0^{\infty} e^{itx} dF(x) = E(e^{itx}),$$

en supposant que la variable aléatoire prend toute valeur entre 0 et ∞ . Pour les lois absolument continues, de densité $f(x)$, on a

$$(12) \quad \varphi(t) = \int_0^{\infty} e^{itx} f(x) dx.$$

Dans notre cas, qui est celui des probabilités discontinues,

$$(13) \quad \varphi(t) = \sum_{m=0}^n e^{itm} P_m.$$

Cette fonction caractéristique est un polynôme de degré n en $u = e^{it}$, on peut donc la développer suivant les puissances de

$$v = e^{it} - 1 = u - 1.$$

On aura

$$(14) \quad \varphi(t) = \sum_{k=0}^n \frac{v^k}{k!} [M_k] = \Phi(v),$$

M_k étant les moments factoriels :

$$(15) \quad [M_k] = E[m(m-1) \dots (m-k+1)].$$

Dans le cas des tirages contagieux,

$$(16) \quad [M_k] = n(n-1) \dots (n-k+1) \frac{p(p+\gamma) \dots [p+(k-1)\gamma]}{1(1+\gamma) \dots [1+(k-1)\gamma]}$$

et

$$\Phi(v) = \sum_{k=0}^n v^k C_n^k \frac{\frac{p}{\gamma} \left(\frac{p}{\gamma} + 1\right) \dots \left(\frac{p}{\gamma} + k - 1\right)}{\frac{1}{\gamma} \left(\frac{1}{\gamma} + 1\right) \dots \left(\frac{1}{\gamma} + k - 1\right)},$$

ou

$$(17) \quad \Phi(v) = \sum_{k=0}^n v^k C_n^k \frac{\Gamma\left(\frac{p}{\gamma} + k\right) \Gamma\left(\frac{1}{\gamma}\right)}{\Gamma\left(\frac{p}{\gamma}\right) \Gamma\left(\frac{1}{\gamma} + k\right)}.$$

Si l'on introduit l'intégrale eulérienne de première espèce,

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx = \frac{\Gamma(\alpha) \Gamma(\beta)}{\Gamma(\alpha + \beta)},$$

on aura, comme il est facile à voir,

$$(18) \quad \Phi(v) = \frac{1}{B\left(\frac{p}{\gamma}, \frac{q}{\gamma}\right)} \int_0^1 x^{\frac{p}{\gamma}-1} (1-x)^{\frac{q}{\gamma}-1} (1+vx)^n dx.$$

Dans notre cas,

$$\frac{p}{\gamma} = \frac{1}{r}, \quad \frac{q}{\gamma} = \frac{1-p}{\gamma} = \frac{1}{\gamma} - \frac{1}{r}.$$

Si l'on écrit encore

$$\frac{\gamma x}{\mu r} = \xi$$

on retrouve le calcul fait en a).

Remarquons que si $n \rightarrow \infty$,

$$(1+vx)^n = \left[1 + \left(e^{i\frac{\mu r}{n\gamma}} - 1\right) \frac{\gamma \xi}{\mu r}\right]^n \rightarrow e^{i\frac{\mu r}{\gamma} \xi},$$

donc

$$(19) \quad \Phi(v) \rightarrow \frac{1}{(\mu r)^{\frac{1}{r}} \Gamma\left(\frac{1}{r}\right)} \int_0^{\frac{\mu r}{\gamma}} \xi^{\frac{1}{r}-1} e^{-f_\gamma(\xi) + o(v)} e^{i\frac{\mu r}{\gamma} \xi} d\xi.$$

Faisons encore tendre γ vers 0, $\frac{1}{\gamma} \rightarrow \infty$, $f_\gamma(\xi) \rightarrow \frac{\xi}{\mu r} = f_0(\xi)$, et

$$(20) \quad \lim \Phi(v) = \int_0^1 \frac{1}{(\mu r)^{\frac{1}{r}} \Gamma\left(\frac{1}{r}\right)} \xi^{\frac{1}{r}-1} e^{-f_0(\xi)} e^{i\frac{\mu r}{\gamma} \xi} d\xi.$$

En tenant compte de la formule (12) on voit donc que $\varphi(t)$ tend vers la fonction caractéristique de la loi dont la densité a pour expression

$$(21) \quad \frac{1}{(\mu r)^{\frac{1}{r}} \Gamma\left(\frac{1}{r}\right)} \xi^{\frac{1}{r}-1} e^{-f_0(\xi)}$$

et c'est ainsi la densité de la loi-limite de P_m .

Si

$$x = \left(t + \frac{g}{t^{r-1}} \right)^r,$$

on aura

$$f_\gamma(x) = f_\gamma(t^r) + \frac{g}{\mu} + O\left(\frac{1}{t}\right)$$

et

$$x^{\frac{1}{r}-1} = t^{1-r} \left\{ 1 + O\left(\frac{1}{t^{2r-1}}\right) \right\},$$

c'est-à-dire

$$(22) \quad P_m = \frac{\mu r}{\gamma n} \frac{e^{-f_\gamma(t^r) + o(\gamma) + o\left(\frac{1}{t}\right)}}{(\mu r)^{\frac{1}{r}} \Gamma\left(\frac{1}{r}\right) t^{r-1}} e^{-\frac{g}{\mu}} \left\{ 1 + O\left(\frac{1}{t^{2r-1}}\right) \right\}.$$

La probabilité de l'inégalité

$$\frac{n\gamma}{\mu r} \left(t + \frac{a}{t^{r-1}} \right)^r < m < \frac{n\gamma}{\mu r} \left(t + \frac{b}{t^{r-1}} \right)^r$$

s'exprime alors par

$$\begin{aligned} P(t_1, t_2) &= \sum_{\frac{n\gamma}{\mu r} \left(t + \frac{a}{t^{r-1}} \right)^r < m < \frac{n\gamma}{\mu r} \left(t + \frac{b}{t^{r-1}} \right)^r} P_m = \\ &= \frac{\mu r}{n\gamma} \frac{e^{-f_\gamma(t^r) + o(\gamma) + o\left(\frac{1}{t}\right)}}{(\mu r)^{\frac{1}{r}} \Gamma\left(\frac{1}{r}\right) t^{r-1}} \left\{ 1 + O\left(\frac{1}{t^{2r-1}}\right) \right\} \sum_{g=a}^b e^{-\frac{g}{\mu}}. \end{aligned}$$

Si l'on remplace m par $m+1$, on peut admettre, en première approximation, que g augmente de $\frac{\mu}{n\gamma}$, donc

$$\sum_{g=a}^b e^{-\frac{g}{\mu}} = \frac{n\gamma}{\mu} \int_a^b e^{-\frac{s}{\mu}} ds = \frac{n\gamma}{\mu r} \mu r \left(e^{-\frac{a}{\mu}} - e^{-\frac{b}{\mu}} \right).$$

Nous avons enfin

$$(23) \quad P(t_1, t_2) = \frac{e^{-t_1 \gamma^{(r)}}}{t_1^{r-1}} \frac{(\mu r)^{1-\frac{1}{r}}}{\Gamma\left(\frac{1}{r}\right)} \left(e^{-\frac{a}{\mu}} - e^{-\frac{b}{\mu}} \right) \{1 + O(1)\}.$$

On voit l'analogie entre ce problème et celui traité dans 2. Un changement de variable opéré sur (9) permettrait d'ailleurs de trouver l'expression (4').

La formule (23) reste valable pour $a=0$, $b=+\infty$, comme on le montre d'une façon tout-à-fait analogue à celle de M. KHINTCHINE. Donc

$$(24) \quad P(t) = \frac{e^{-t \gamma^{(r)}}}{t^{r-1}} \frac{(\mu r)^{1-\frac{1}{r}}}{\Gamma\left(\frac{1}{r}\right)} \{1 + O(1)\},$$

et alors

$$(25) \quad \frac{P(t_1, t_2)}{P(t)} = e^{-\frac{a}{\mu}} - e^{-\frac{b}{\mu}}.$$

Il faut encore voir la signification de μ .

Il est facile à montrer que

$$E(m) = \sum_{m=0}^n m P_m = np = \frac{n\gamma}{r},$$

donc

$$\mu = E\left(\frac{\mu r}{n\gamma} m\right) = E(x).$$

Le résultat de ce problème particulier n'est donc pas essentiellement différent de celui obtenu pour la loi de LAPLACE, si on fait les recherches au même point de vue.

(Reçu le 26 juillet 1935)