Quelques idées au sujet du mémoire de M. G. Neymann „L'estimation statistique traitée comme un problème classique de probabilité".[1] [2]

Par L. KALOUJNINE à Paris.

Soit T l'espace des résultats possibles d'une certaine expérience dont on sait que sa loi de distribution w fait partie d'un certain ensemble de lois de distribution W. On notera par t les éléments de T. ($T = \{t\}$).

Soit $F(t)$ une application de T dans l'ensemble des sous-ensembles non-vides de W. (C'est à dire, 2^W désignant l'ensemble des sous-ensembles de W, on a pour tout $t \in T$ $F(t) \in 2^W$ et $F(t) \neq \emptyset$). Pour tout $w \in W$ soit $F^{-1}(w)$ l'ensemble de tous les $t' \in T$ tels que $F(t')$ couvre w. (C'est à dire $F^{-1}(w)$ est l'ensemble de tous les $t' \in T$ tels que $w \in F(t')$). On exigera en outre que si t parcourt T, $F(t)$ couvre W, ou, autrement dit, pour tout $w \in W$ il existe un $t \in T$, tel qu'on ait $w \in F(t)$.

Cela étant, désignons pour tout w par $P_w(F^{-1}(w))$ la probabilité (dont on supposera l'existence) de l'ensemble $F^{-1}(w)$ sous l'hypothèse que la loi de distribution w a lieu. On posera

$$\alpha(F) = \underset{w \in W}{\text{borne inférieure}} \, (P_w(F^{-1}(w))).$$

L'application F ainsi définie *sera appellée une estimation relative à l'espace d'expériences T et à l'ensemble des hypothèses W.* $\alpha(F)$ sera dit la *pseudo-probabilité de l'estimation F.* Dans le cas où $0 < \alpha(F) < 1$ on parlera *d'une estimation* non-banale.

L'estimation ainsi définie est une règle de conduite qui se résume comme suit: *toute fois que le résultat d'une expérience est un $\bar{t} \in T$ je declarerai* "*La*

[1] Colloque consacré à la théorie des probabilité (Genéve 1937). — *Act. Scient. et Ind.* **739.** p. 25—57.

[2] Les idées exposées dans la présente Note ne sont certainement pas nouvelles. Ils ont été formulées sous des aspects différents par plusieurs auteurs, notement par G. NEYMANN, E. S. PEARSON, A. WALD, H. CRAMÉR et bien d'autres. J'espère que mon présent exposé et surtout l'exemple du "jeu d'estimation" contribuera à une clarification de la notion d'estimation si importante dans la statistique mathématique.

loi de distribution w qui a lieu est un élément de $F(\bar{t})$". Comme on verra ci-dessous *on ne peut pas* affirmer qu'en suivant cette règle de conduite la fréquence des déclarations justes tendra en probabilité vers $\alpha(F)$ (en général cette fréquence *ne convergera pas* en probabilité). Par contre, il est possible de formuler une autre propriété de cette règle de conduite, qui pratiquement rend les mêmes services. On peut affirmer, en effet, qu'en appliquant l'estimation F un grand nombre de fois la fréquence de jugements justes sera à la longue $\geqq \alpha(F)$. Plus précisement a lieu "la loi des grands nombres" suivante:

Loi des grands nombres: F étant une estimation non banale relative à T et à W et $\alpha(F)$ étant sa pseudo-probabilité, pour tout $\varepsilon < \alpha(F)$ et pour tout $\delta > 0$ il existe un entier M, tel que si l'on applique l'estimation $m > M$ fois, on a, en désignant par μ le nombre de jugements justes,

$$P\left(\frac{\mu}{m} > \varepsilon\right) > 1 - \delta.$$

Je donne maintenant un exemple très simple d'une estimation et je tacherai d'y préciser le sens des notions introduites ainsi que l'importance de "la loi des grands nombres" signalée. La généralisation de la discussion aux cas plus généraux est immédiate. On verra également que la démonstration de "la loi des grands nombres" est fort simple, car elle se réduit à celle relative au schema de Poisson (schema de Bernoulli généralisé).

Voici l'exemple en question: soit donnée une urne contenant deux boules, dont on sait qu'elles ne peuvent être que blanches ou noires. Il y a donc trois hypothèses: $w_1 = \{\circ\circ\}$, $w_2 = \{\bullet\bullet\}$ et $w_3 = \{\circ\bullet\}$. L'expérience consiste en n tirages au hasard successifs avec remplacement de la boule. 2^n résultats d'expérience sont possibles qui constituent dans notre cas l'espace T.

L' estimation $F(t)$ sera définie de la manière suivante: si t_1 (resp. t_2) est l'expérience qui a produit n boules blanches (resp. noires) on posera

$$F(t_1) = w_1,\ F(t_2) = w_2 \qquad \text{et pour } t' \neq t_1, t_2, \qquad F(t') = w_3.$$

On a $P_{w_1}(F^{-1}(w_1)) = P_{w_2}(F^{-1})(w_2)) = 1$ et $P_{w_3}(F^{-1})w_3)) = 1 - \frac{1}{2^{n-1}}$

donc $\alpha(F) = 1 - \frac{1}{2^{n-1}}$.

Pour mettre en relief "la loi des grands nombres" nous imaginezons le jeu suivant qu'on appellera "*jeu d'estimation*": A joue avec B. Au jeu participe également un troisième joueur C (joueur impartial) qui n'est interessé ni au gains de A ni à ceux de B. C met deux boules dans l'urne w_1, ou bien w_2, ou bien w_3. A effectue n tirages au hasard avec remplacement de la boule et applique au résultat de l'expérience l'estimation F définie ci-dessus. Ensuite l'urne est ouverte. Si la déclaration de A était juste, il a gagné, sinon B.

Ce jeu étant repeté un grand nombre de fois, *que peut on dire sur la fréquence des gains de A?*

[Si C choisit w_1, w_2, w_3 au hasard de manière que la probabilité de w_ν soit γ_ν, alors dans chaque jeu la probabilité de gain de A est $q = \Sigma_\nu \gamma_\nu P_{w_\nu}(F^{-1}(w_\nu))$ (schema de Bayes) et la fréquence de gains de A tendra en probabilité vers $q \geqq \alpha(F)$.] Supposons que C adopte un système en vertu duquel dans le μ-ième jeu il met dans l'urne $\varphi(\mu)$, où $\varphi(\mu)$ est une des configurations w_1, w_2 où w_3. Ainsi, $\varphi(\mu)$ est une fonction définie sur les entiers positifs et prenant comme valeurs une des "hypothèses" w_1, w_2 w_3 Quelle que soit le système $\varphi(\mu)$ adopté par C on se trouve dans le cas du schema de Poisson. En effet, en désignant par 1 le gain de A, par 0 celui de B, le résultat du μ-ième jeu peut être envisagé comme une variable aléatoire X_μ ne prenant que les valeurs 1 et 0 avec les probabilités respectives $P_{\varphi(\mu)}(F^{-1}(\varphi(\mu)))$ et $1 - P_{\varphi(\mu)}(F^{-1}(\varphi(\mu)))$ Dès lors pour tout $\delta > 0$ il existe un entier M tel que pour $m > M$ on a

$$P\left(\left| \frac{\sum_{\mu=1}^{m} X_\mu}{m} - \frac{\sum_{\mu=1}^{m} P_{\varphi(\mu)}(F^{-1}(\varphi(\mu)))}{m} \right| < \delta \right) > 1 - \delta.$$

Comme d'autre part l'écart quadratique moyen de X_μ ne dépasse pas $\frac{1}{2}$ l'entier M peut être choisi indépendamment du système $\varphi(\mu)$ adopté.

On a toujours $\dfrac{\sum_{\mu=1}^{m} P_{\varphi(\mu)}(F^{-1}(\varphi(\mu)))}{m} \geqq \alpha(F)$, d'où résulte la validité de "la loi des grands nombres" sous la forme signalée plus haut. [Il est évident que si C adopte un système $\varphi(\mu)$ tel que $\dfrac{\sum_{\mu=1}^{m} P_{\varphi(\mu)}(F^{-1}(\varphi(\mu)))}{m}$ ne converge pas avec m, la fréquence des gains de A ne convergera pas en probabilité.]

Il est également possible de démontrer une "loi des grands nombres forte" qui affirme que sous les mêmes hypothèses que pour "la loi des grands nombres" formulée plus haut on a

$$P\left(\lim \inf \frac{\mu}{m} \geqq \alpha(F) \right) = 1.$$

(Reçu le 15 novembre 1949.)