Extrait du cours TOPOMETRIE GENERALE
par J.J. LEVALLOIS Ingénieur en chef Géographe
Institut Géographique National
Edition 1960

ANNEXE

LE CALCUL DES PROBABILITÉS ET LA THÉORIE DES ERREURS

Le but de cette annexe est de rappeler quelques notions très élémentaires du calcul des probabilités et de les utiliser pour préciser mathématiquement les résultats obtenus par voie empirique au chapitre II.

Par cette voie, on peut donner l'équation de la courbe des fréquences, préciser les rapports entre les erreurs moyennes quadratiques, arithmétique et probables, et définir un certain nombre de points que l'étude empirique laisse dans l'ombre.

L'introduction du calcul des probabilités dans la théorie des erreurs accidentelles a soulevé des controverses fort longues : le physicien LIPPMAN avait coutume de dire que les physiciens acceptent la loi générale de distribution des erreurs accidentelles comme une vérité établie par les mathématiciens, et que les mathématiciens la considèrent comme une donnée expérimentale éprouvée par les physiciens. La boutade est célèbre et reflète bien l'aspect de la question. Toutefois les deux points de vue sont conciliables : l'expérience met hors de doute l'existence d'une loi générale de répartition des erreurs accidentelles, suivie assez fidèlement dans le domaine observable ; le calcul des probabilités en donne l'expression asymptotique et permet en outre de déceler si les écarts entre l'expérience et la théorie sont admissibles ou anormaux.

Dans cet exposé nous ferons d'abord un rappel succint des théorèmes fondamentaux du calcul des probabilités. Nous les utiliserons pour étudier :

1° la théorie de la moyenne ;

2° la répartition théorique des erreurs vraies et des écarts ;

3° certaines questions dont nous n'avons pas parlé encore.

Probabilité.

La notion d'égales probabilités est une notion intuitive ; deux événements sont également probables lorsque l'on n'a aucune raison de prévoir l'arrivée de l'un plutôt que celle de l'autre : pile ou face, pair ou impair, tirage au loto, etc...

On appelle probabilité d'un événement, le rapport du nombre des cas favorables à la production de cet événement au nombre total des éventualités, chacune d'entre elles étant considérée comme également probable. La probabilité est donc un nombre positif compris entre 0 et 1 ; ces deux cas représentant l'un la certitude et l'autre l'impossibilité.

Deux événements sont *exclusifs* lorsque la production de l'un quelconque d'entre eux est incompatible avec celle de l'autre. Ils sont *indépendants* si la production de l'un quelconque d'entre eux est indifférente à la production de l'autre.

Si par exemple on jette un seul dé de zanzibar, l'apparition de la face 5 et celle de la face 6 sont deux événements exclusifs.

Lorsque l'on jette (simultanément ou successivements) deux dés identiques, l'apparition d'une face quelconque sur le premier dé et sur le second sont deux événements indépendants.

Le calcul des probabilités envisage fréquemment des éventualités complexes constituées par l'arrivée de plusieurs événements exclusifs ou indépendants. On énonce à ce sujet les deux théorèmes suivants :

Théorème des probabilités totales.

La probabilité de la production de l'un *ou* l'autre de deux événements exclusifs, A ou B, est égale à la somme de leurs probabilités respectives.

Soit en effet N le nombre total des éventualités, a le nombre des cas produisant l'événement A, b celui qui correspond à la production de B. L'événement gagnant (A *ou* B) peut se produire dans $(a+b)$ cas favorables, sa probabilité est donc :

$$P = \frac{a+b}{N} = \frac{a}{N} + \frac{b}{N} = P(A) + P(B).$$

Le théorème se généralise immédiatement à un nombre quelconque d'événements exclusifs.

Remarque : Il peut arriver que l'exclusion ne soit pas totale et que dans c cas par exemple les événements A et B ne soient pas incompatibles. La probabilité totale est alors égale à la somme des probabilités totales respectives diminuée de celle des cas non exclusifs.

En effet à l'événement A concourent a cas favorables

« B « b «

« (A + B) « c «

les c cas sont compatibles avec a et b.

Classons les événements favorables en 3 catégories exclusives :

A apparaît dans $(a - c)$ cas ... $P_1 = \frac{a - c}{N}$

B apparaît dans $(b - c)$ cas ... $P_2 = \frac{b - c}{N}$

A + B apparaît dans c cas ... $P_3 = \frac{c}{N}$

Les événements ainsi dénombrés sont exclusifs donc :

$$P(A \text{ ou } B) = P_1 + P_2 + P_3$$

c'est-à-dire :

$$P = \frac{a + b - c}{N} = P(A) + P(B) - P(A + B).$$

Exemple : Jetons simultanément deux dés de zanzibar, quelle est la probabilité d'obtenir au moins une face 3.

Au premier dé, $P(3) = \frac{1}{6}$.

Au second dé, $P(3) = \frac{1}{6}$.

En outre les deux cas ne sont pas exclusifs puisqu'il y a possibilité d'obtenir le 3 simultanément au premier et au deuxième dé. Cette probabilité est égale à $\frac{1}{36}$; on a donc :

$$P = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}.$$

Lorsque le nombre des éventualités n'est susceptible de produire que deux événements exclusifs A et B, les probabilités de A ou B sont complémentaires à 1 puisque l'on est certain que A ou B se produira :

$$1 - P(A) = P(B).$$

Reprenons l'exemple précédent et cherchons la probabilité de n'obtenir aucune face 3. Nous avons (en anticipant sur le théorème 2) :

$$P = \left(\frac{5}{6}\right) \cdot \left(\frac{5}{6}\right) = \frac{25}{36}.$$

$\frac{5}{6}$ représente en effet pour l'un des dés, la probabilité de n'obtenir aucune face 3; comme le jet du deuxième dé est indépendant de celui du premier, la probabilité composée de l'événement négatif est égale à $\frac{5}{6} \times \frac{5}{6}$.

La probabilité de sortir le 3 au moins une fois est donc :

$$1 - \frac{25}{36} = \frac{11}{36}.$$

Théorème des probabilités composées.

La probabilité de production de *l'ensemble* de deux événements indépendants A *et* B est égale au produit de leurs probabilités respectives.

Soit a le nombre des cas favorables à la production de A, b celui des cas favorables à l'événement B. On a :

$$P(A) = \frac{a}{M} \qquad P(B) = \frac{b}{N}$$

Le nombre total des cas possibles est égal à M × N, puisque, par suite de l'indépendance, à toute éventualité M peuvent être associées chacune des éventualités N.

De même le nombre des cas favorables à l'événement (A + B) est $a \times b$, puisque l'événement gagnant ne peut être réalisé qu'en associant l'un quelconque des a cas favorables à A avec l'un quelconque des b cas favorables à B. Par suite :

$$P(A + B) = \frac{a \times b}{M \times N} = \frac{a}{M} \times \frac{b}{N} = P(A) \times P(B).$$

C'est bien ce résultat que nous avons appliqué dans l'exemple précédent pour calculer la probabilité composée de n'obtenir aucune face 3.

Principe ou théorème de Bayes.

Soient n événements $A_1A_2 \dots A_n$ de probabilités respectives $a_1a_2 \dots a_n$ et supposons que nous sachions à posteriori que c'est l'un des p premiers qui s'est produit.

On peut se demander quelle est la probabilité pour que ce soit A_i.

Des événements qui avaient a priori mêmes probabilités de se produire les ont conservées après production, autrement dit, nous avons autant de chances pour que ce soit l'un ou l'autre; il en résulte donc qu'a posteriori, les événements A ont conservé des probabilités proportionnelles à leurs probabilités a priori et, puisque $\Sigma_1^p a'_i = 1$ (a posteriori), on a :

$$P(A_i) = \frac{a_i}{a_1 + a_2 + \dots + a_p}.$$

Soit alors B un événement dont l'arrivée est conditionnée par la production antérieure d'événements $A_1 A_2 \ldots A_n$.

L'événement B avait dans ces différentes éventualités des probabilités $b_1\ b_2 \ldots b_n$. La probabilité totale de l'événement B est alors :

$$b = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n.$$

Si l'on se demande quel est celui de tous les événements A qui a déclenché la production de B, la probabilité pour que soit l'événement A_i sera donc :

$$P(A_1) = \frac{a_i b_i}{b}$$

C'est la probabilité a posteriori de l'événement A_i en tant que cause des événements B.

Les 3 théorèmes précédents sont les théorèmes fondamentaux du calcul élémentaire des probabilités ; nous allons les appliquer à quelques exemples tout à fait classiques de manière à en illustrer les modalités d'application.

Exemples d'applications : 1° Dans une donne de bridge, quelle est la probabilité pour qu'un joueur ait les 4 as dans son jeu.

Le nombre de cas possibles est celui des donnes, c'est le nombre des combinaisons de 52 cartes 13 à 13.

Pour avoir le nombre des cas favorables, supposons que le joueur ait obtenu les 4 as d'entrée. Il complétera son jeu avec 9 cartes prises dans le lot des 48 qui restent. La probabilité cherchée est donc égale à :

$$\frac{C_{48}^{9}}{C_{52}^{13}} = \frac{48\,!}{9\,!\,40\,!} \times \frac{13\,!\,39\,!}{52\,!}$$

2° On lance 5 dés ordinaires, quelle est la probabilité d'avoir en un coup sec, une paire, un brelan, un carré, 5 faces identiques, étant entendu que l'on ne veut pas tenir compte des points supérieurs à celui qui est demandé.

Le raisonnement est identique dans les 4 cas. Fixons-nous une face en demande, par exemple le 3.

Supposons que nous désirions sortir p faces 3 et p seulement et ceci sur les p premiers dés lancés.

Nous avons les probabilités suivantes :

sortir la face 3 $\frac{1}{6}$ au 1er dé

sortir la face 3 $\frac{1}{6}$ au p^{me} dé

ne ~~pas~~ la sortir aux autres $\frac{5}{6}$ au $p+1^{me}$ dé, etc.

La combinaison cherchée peut être obtenue avec p quelconques des 5 dés jetés, et peut donc s'obtenir de C_5^p manières différentes, toutes exclusives. Les probabilités cherchées sont donc de la forme :

$$C_5^p \times \left(\frac{1}{6}\right)^p \times \left(\frac{5}{6}\right)^{5-p}$$

ce qui donne pour une paire : $C_5^2 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3$

un brelan : $C_5^3 \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^2$

un carré : $C_5^4 \left(\frac{1}{6}\right)^4 \cdot \left(\frac{5}{6}\right)^1$

5 faces : $\left(\frac{1}{6}\right)^5$

Il va de soi que si l'on ne spécifie pas la nature de la face que l'on désire amener, ces probabilités doivent être multipliées par 6.

Nous allons maintenant étudier un problème de probabilités continues, c'est-à-dire où le dénombrement des cas ne peut être entrepris par comptage.

Problème de l'aiguille.

On jette une aiguille sur une feuille horizontale de papier sur laquelle sont tracées des lignes parallèles équidistantes. Quelle est la probabilité pour que l'aiguille touche l'une des lignes. On suppose que la longueur de l'aiguille $2l$ est inférieure à l'intervalle des parallèles $2a$.

Soit G le centre de gravité de l'aiguille, on remarque que le point de chute de G dans le sens des parallèles est indifférent. Pour que l'aiguille touche une ligne il faut :

1° que G tombant à une distance x de la ligne la plus proche, la distance G soit inférieure à a, c'est-à-dire que le cercle de centre G de rayon l coupe cette ligne ;

2° que l'aiguille tombe dans l'angle 2α tel que $2\alpha = \mathrm{IGI'}$ découpé dans ce cercle par la droite rencontrée.

Soit alors x la distance HG. Pour que G tombe dans un intervalle x, $x + dx$ de la distance HH' on a la probabilité $\frac{dx}{2a}$.

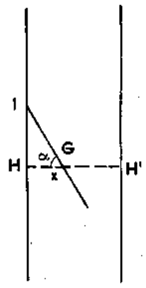

Fig. 112

Pour que l'aiguille touche la ligne, sa direction devra être comprise dans l'angle 2α, au centre du cercle G. Cette probabilité est égale par conséquent à $\frac{2\alpha}{2\pi}$. La probabilité pour que l'on ait à la fois G dans l'intervalle x, $x + dx$ et l'aiguille dans la bonne direction est donc une probabilité composée :

$$P = \frac{\alpha}{\pi} \times \frac{dx}{2a}.$$

La probabilité totale de toucher la ligne AA' sera donc Σ (P), il est à remarquer qu'elle est nulle si le cercle ne peut couper la droite.

D'autre part $\frac{x}{l} = \cos\alpha$ d'où $dx = - l \sin\alpha \, . \, d\alpha$:

$$P(x)\, dx = \frac{\alpha}{\pi} \times \frac{-l \sin\alpha}{2a}\, d\alpha$$

$$P = \int_{\frac{\pi}{2}}^{0} \frac{-\alpha}{\pi} \times \sin\alpha \, d\alpha \times \frac{l}{2a}$$

$$= \frac{l}{2\pi a}\left[\left(-\alpha\cos\alpha\right)_{\frac{\pi}{2}}^{0} + \int_{\frac{\pi}{2}}^{0} \cos\alpha\, d\alpha\right] = -\frac{l}{2\pi a}$$

Comme la figure est symétrique et puisque l'aiguille a deux extrémités et qu'il y a exclusion entre contact d'une ligne ou de l'autre, la probabilité cherchée est finalement :

$$P = \frac{2l}{\pi a}$$

Elle dépend du nombre π... On a exécuté pratiquement cette expérience pour savoir si elle s'accordait avec la théorie. Le résultat constitue une éclatante vérification de celle-ci ; sur un nombre de 5.000 épreuves on a calculé pour π la valeur 3,15. Il est difficile d'obtenir mieux dans la vérification des lois physiques : avec une aiguille de 36^{mm} et un écart de lignes de 45^{mm}, Wolf a obtenu : 2.532 sécances au lieu de 2.546 calculées :

$$\frac{1}{\pi} = 0.3169 \text{ au lieu de } 0.3183.$$

Lois de probabilité (1).

Une variable est dite éventuelle lorsqu'elle peut prendre l'une quelconque des valeurs $x_1 x_2 \dots x_i \dots x_n$ d'une suite continue ou non de valeurs auxquelles sont attachées les probabilités $a_1 a_2 \dots a_n$. Toutes les valeurs possibles de x sont énumérées dans la suite x_1. Il en résulte que $\Sigma a_i = 1$.

Par définition la valeur probable de x est la moyenne du premier ordre :

$$\frac{\Sigma a_i x_i}{\Sigma a_i} = \Sigma a_i x_i$$

nous la noterons :

$$m_1(x) = \Sigma a_i x_i.$$

La moyenne d'ordre p de la variable x sera par définition la valeur m_p telle :

$$[m_p(x)]^p = \Sigma a_i x_i^p$$

L'erreur commise est la valeur $x - m_1(x)$, et l'erreur moyenne quadratique :

$$\eta^2(x) = \Sigma a_i [x - m_1(x)]^2$$

On remarque immédiatement que la valeur moyenne de l'erreur $x - m_1(x)$ est nulle puisque cette valeur moyenne est :

$$\Sigma a_i [x_i - m_1(x)] = m_1(x) - m_1(x).$$

En outre on a en développant l'erreur moyenne quadratique :

$$\eta^2(x) = \Sigma a_i x_i^2 - 2\Sigma a_i x_i \, . \, m_1(x) + \Sigma a_i m_1(x)^2$$
$$\Sigma a_i = 1$$
$$\eta^2(x) = m_2^2(x) - m_1^2(x).$$

Composition des lois de probabilités.

Soient deux variables éventuelles x et y. Considérons la variable éventuelle $z = x + y$, les lois de probabilités de x et y étant quelconques.

Nous nous proposons de chercher la valeur probable de z et son erreur moyenne quadratique.

La probabilité de l'événement z_{ij} composé de $z_{ij} = x_i + y_j$ *est* alors $a_i b_j$ et la valeur probable $m_1(z)$ est donnée par sa définition :

$$m_1(z) = \Sigma a_i b_j (x_i + y_j)$$
$$= \Sigma b_j \Sigma a_i x_i + \Sigma a_i \Sigma b_j y_j.$$
$$m_1(z) = m_1(x) + m_1(y)$$

La valeur moyenne d'une somme de variables éventuelles est la somme des valeurs moyennes de ces variables.

(1) Toute cette partie du Cours suit de très près l'exposé de M. P. Lévy dans son « Cours d'Analyse de l'Ecole Polytechnique », Gauthier-Villars, 1931.

Cherchons la moyenne d'ordre deux de z.

$$m_2^2(z) = \Sigma a_i b_j (x_i + y_j)^2 = \Sigma a_i x_i^2 \Sigma b_j + 2\Sigma a_i x_i \Sigma b_j y_j + \Sigma a_i \Sigma b_j y_j^2$$

$$m_2^2(z) = m_2^2(x) + 2m_1(x)\, m_1(y) + m_2^2(y)$$

Si les termes tels que $m_1(x)$, $m_1(y)$ sont nuls, ce qui est précisément le cas pour les valeurs $x - m_1(x)$, $y - m_1(y)$, on voit que le terme rectangle disparaît. On a donc :

$$\eta^2(z) = \eta^2(x) + \eta^2(y)$$

Quelles que soient les lois de probabilités auxquelles sont astreintes les variables éventuelles x et y. L'erreur moyenne quadratique d'un somme est égale à la racine carrée de la somme des carrés des erreurs moyennes quadratiques composantes. En particulier, si les lois de probabilités sont les mêmes pour les variables éventuelles composant z, on aura :

$$\eta(z) = \eta\sqrt{n}$$

n étant le nombre des variables composant z.

Inégalité fondamentale.

Considérons une variable éventuelle quelconque z. Son erreur moyenne quadratique $\eta(z)$ est telle que :

$$\eta^2(z) = \Sigma a_i [z_i - m_1(z)]^2.$$

Cette quantité donne une idée approchée de la valeur de l'erreur commise sur z lorsque l'on choisit $m_1(z)$ comme valeur approchée de z — il suffit de regarder sa définition pour s'en convaincre.

Remarquons que des erreurs très *supérieures* à $\eta(z)$ sont très peu probables. Supposons en effet que z_j soit une valeur de z telle que $z_j - m_1(z) = k\,.\,\eta(z)$, k étant un nombre plus grand que un.

Soit C la probabilité de $z_j - m_1(z)$. Nous pouvons écrire.

$$\eta^2(z) = \Sigma a'_i [z_i - m_1(z)]^2 + C\,[z_j - m_1(z)]^2 = \Sigma a'_i [z_i - m_1(z)]^2 + C\,.\,k^2\eta^2(z).$$

Cette expression montre que $C < \dfrac{1}{k^2}$ Puisque tous les termes sont positifs. Cette inégalité très importante est appelée parfois inégalité de Bienaymé.

Théorème de Bernouilli ou Loi des grands nombres.

La fréquence d'un événement A tend vers sa probabilité lorsque le nombre des épreuves devient très grand.

Nous avons vu que dans la composition des lois de probabilités identiques, l'erreur moyenne quadratique de la résultante était liée à celle des composantes par la relation $\eta(z) = \eta\sqrt{n}$, nous avons vu d'autre part que les valeurs de l'erreur de z très supérieures à $\eta(z)$ étaient très peu probables. Il en résulte que si nous considérons la variable éventuelle $\dfrac{z}{n} = \dfrac{x + y + \dots}{n}$ l'intervalle de variation de $\dfrac{z}{n} - m_1(z)$ sera de l'ordre de $\dfrac{1}{\sqrt{n}}$ et tendra vers zéro avec n, ou du moins que des valeurs très discordantes de cette différence seront très peu probables. Considérons alors un événement A pouvant se produire concurremment à d'autres événements B dans diverses hypothèses $H_1 H_2 \dots H_n$ et soient $a_1 a_2 \dots a_n$ les probabilités de A dans les diverses hypothèses considérées.

Associons à l'événement A une variable éventuelle x à laquelle nous attacherons la valeur 1 si A se réalise et la valeur 0 dans le cas contraire.

Le nombre des réalisations de A est alors au bout de n coups :

$$z = x_1 + x_2 + \dots x_n$$

la fréquence est $\frac{z}{n}$.

Les probabilités de production de A sont alors :

$$a_1 a_2 \dots a_n$$

associées à $x_1 = 1$; celles de non productions sont :

$$1 - a_1 \qquad 1 - a_2 \qquad \dots \qquad 1 - a_n$$

associées à $x_j = 0$; donc :

$$\eta^2(z) = m_2^2 - m_1^2 = \sum_1^n a_i(1 - a_i) < \frac{1}{4}\, n.$$

La valeur moyenne de z est :

$$m_1(z) = \Sigma a_i x_i + \Sigma(1 - a_j)\, x_j = \Sigma a_i.$$

On a donc :

$$m_1(z) - z \sim K\sqrt{n}$$

et :

$$\frac{m_1(z)}{n} - \frac{z}{n} \sim \frac{K}{\sqrt{n}}$$

c'est-à-dire :

$$\frac{a_1 + a_2 + \dots + a_n}{n} - \frac{z}{n} \sim \frac{K}{\sqrt{n}}$$

or la quantité $\frac{z}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$ est la fréquence de l'événement A. La valeur moyenne de la probabilité de A diffère donc de la fréquence d'une quantité infiniment petite avec $\frac{1}{\sqrt{n}}$ et si nous supposons en outre que $a_1 = a_2 = \dots a_n$, on a :

$$a - \frac{z}{n} \sim \sqrt{\frac{a(1 - a)}{n}}$$

où a est la probabilité.

Expression qui tend vers zéro pour n infini et qui nous donne en outre l'écart moyen quadratique entre la fréquence et la probabilité, c'est-à-dire l'écart moyen quadratique de la fréquence.

Ce résultat très important est connu sous le nom de loi des grands nombres — la fréquence tend vers la probabilité —. Ses applications sont multiples, son application est la base de la statistique. Il prouve en particulier que la valeur la plus probable d'une inconnue donnée par des expériences est celle qui correspond à la plus grande fréquence, et tel est le cas de la moyenne arithmétique dans le cas des observations directes. En outre, si nous comparons les valeurs de l'erreur moyenne quadratique ε au sens du Chapitre II et η au sens du calcul des probabilités, nous voyons que ces erreurs sont égales à la limite, puisque la fréquence tend vers la probabilité et que $\Sigma\, v^2$ fait intervenir la fréquence.

LOI DE PROBABILITÉ DES ERREURS ACCIDENTELLES

Les épreuves répétées dans le cas du jeu de pile ou face.

Considérons deux éventualités exclusives dont la probabilité totale soit égale à 1 et supposons par exemple que chaque éventualité isolée soit de probabilité $\frac{1}{2}$ comme dans le cas du jeu de pile ou face.

Nous conviendrons de noter le coup pile par le signe $+$ et le coup face par le signe $-$ Répétons n fois l'expérience et cherchons la probabilité d'obtenir p fois pile et q fois face ($p+q=n$), quelle que soit par ailleurs la répartition interne des signes $+$ et $-$. Pour fixer les idées, si sur une partie de 5 coups nous désirons obtenir deux fois pile et 3 fois face, cette éventualité peut se produire dans les 10 cas suivants :

$++---$	$-+-+-$
$-++--$	$--+-+$
$--++-$	$+--+-$
$---++$	$-+--+$
$+-+--$	$+---+$

et dans ces 10 cas seulement.

Dans le cas général, considérons une éventualité ; composé de p fois pile et q fois face ; cherchons sa probabilité.

Il y a évidemment 2^n cas possibles. Chaque cas favorable ne différera du cas envisagé que par une permutation au moins d'un signe $+$ et d'un signe $-$. On peut ainsi réaliser C_n^p ou C_n^q dispositions différentes.

La probabilité cherchée est donc :

$$\frac{1}{2^n} \times C_n^p = \frac{1}{2^n} \times C_n^q$$

Les probabilités de distribution des combinaisons possibles varient donc comme les coefficients du binôme de NEWTON. On peut alors dresser pour le cas où $n = 10$ par exemple, le tableau ci-dessous :

$+$	$-$	Probabilités	Valeur numérique
10	0	$\frac{1}{1024}$	0.0010
9	1	$\frac{10}{1024}$	0.0098
8	2	$\frac{45}{1024}$	0.0439
7	3	$\frac{120}{1024}$	0.1172
6	4	$\frac{210}{1024}$	0.2051
5	5	$\frac{252}{1024}$	0.2461
$-$	$+$	**Probabilités**	**Valeur numérique**

On voit que l'éventualité 5 — 5 est axe de symétrie et qu'en conséquence choisissant cette éventualité comme origine, portant en abscisse le numéro de la combinaison on peut construire la courbe admettant en ordonnée la probabilité d'une combinaison donnée. En joignant les différents sommets par un trait continu, cette courbe (figure 113) apparaît au premier coup d'œil comme une stylisation des courbes empiriques déjà étudiées précédemment. Cette parenté n'est pas fortuite : une erreur accidentelle est de la forme $e = \pm \varepsilon_1 \pm \varepsilon_2 \pm \ldots$ les ε étant du même ordre de grandeur. La valeur de e ne dépend donc en définitive que de l'alternance des signes $+$ et $-$ et pour des mesures de même précision n et ε seront les mêmes d'une expérience à l'autre.

La distribution des erreurs accidentelles doit donc coïncider sensiblement avec la distribution limite des coups pile et face répartis en séries d'un nombre infiniment grand de coups. La distribution de ces erreurs doit être du type de $\frac{1}{2^n} C_n^p$ lorsque n tend vers l'infini. Nous allons étudier cette limite.

Théorème de Bernouilli.

Lorsque le nombre des épreuves répétées devient infini :

1°) la combinaison la plus probable est celle qui comporte autant de coups pile que de coups face ;

2°) les points représentatifs des coefficients tendent vers les points de la courbe

$$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Supposons en effet n très grand et p variable entre 0 et n.

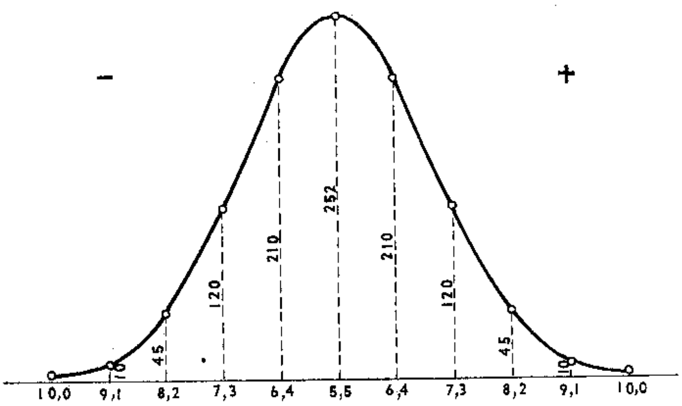

Fig. 113

Le maximum de l'expression $\frac{C_n^p}{2^n}$ sera atteint lorsque le dénominateur $p\,!\,(n-p)\,!$ sera minimum.

Transformons les factorielles par la formule de Stirling : $x\,! \sim \sqrt{2\pi x}\,\frac{x^x}{e^x}$, nous avons :

$$C_n^p = A p^p \,.\, (n-p)^{(n-p)}.$$

Prenons les logarithmes et dérivons, il vient :

$$p \times \frac{1}{p} + \log p - (n-p)\,.\,\frac{1}{n-p} - \log (n-p) = 0$$

c'est-à-dire :

$$\log p = \log (n-p) \qquad\qquad p = \frac{n}{2}.$$

La combinaison la plus probable est donc bien celle qui correspond à la réalisation des piles et faces en nombres égaux et ceci n'est qu'un cas particulier de la loi des grands nombres.

3°) cherchons maintenant la loi de probabilité continue dont dépend la combinaison $p\ q$ si n devient très grand.

Attribuons à chaque coup un gain $X = \pm 1$ suivant le coup amené.

La valeur probable de ce gain est

$$m_1 = \frac{1}{2} \times 1 + \frac{1}{2} \times (-1) = 0$$

$$m_2^2 = \frac{1}{2} + \frac{1}{2} = 1 \qquad\qquad \eta(X) = 1 \text{ puisque } \eta^2 = m_2^2 - m_1^2.$$

Au bout de n coups, la valeur probable du gain est :

$$m_1(X) = 0 \qquad \text{et} \qquad \eta(X) = \sqrt{n}$$

Le gain réel X est égal à :

$$X = m_1(X) + \eta(X) . x$$

x étant ce que nous appellerons l'*erreur réduite,* erreur dont la valeur probable est nulle et la valeur moyenne quadratique égale à 1.

Nous avons alors $X = x\sqrt{n}$, et les valeurs éventuelles du gain sont d'autre part :

$$+n \qquad +(n-2)\ \ldots \qquad (n-2p)\ \ldots \qquad -n$$

On peut écrire :

$$n - 2p = x\sqrt{n}$$

d'où :

$$p = \frac{n}{2} - \frac{x}{2}\sqrt{n}.$$

La probabilité de ce gain est d'autre part :

$$\frac{1}{2^n}\, C_n^p$$

Transformons cette expression par la formule de STIRLING :

$$P(p, q) = \frac{\sqrt{2\pi n}}{\sqrt{2\pi p}} \cdot \frac{n^n}{p^p} \cdot \frac{1}{\sqrt{2\pi(n-p)}\,(n-p)^{n-p}} \cdot \frac{e^{p+n-p}}{e^n}\,\frac{1}{2^n}$$

$$P(p, q) = \frac{\sqrt{n}}{\sqrt{2\pi p(n-p)}}\,\frac{1}{\left(\frac{2p}{n}\right)^p\left(\frac{2(n-p)}{n}\right)^{(n-p)}}.$$

Dans un petit intervalle de variation dp de p, la probabilité totale est égale à :

$$P(p, q)\, dp \qquad \text{avec} \qquad |dp| = \sqrt{\frac{n}{2}}\, dx.$$

D'Où :

$$P(p, q)\, dp = P(x)\, dx.$$

Substituons les valeurs de la variable p en fonction de x.

$$P(x)\,dx = \frac{\sqrt{n}}{\sqrt{2\pi p\,(n-p)}}\,\frac{\sqrt{\frac{n}{2}}\,.\,dx}{\left[\dfrac{2\left(\frac{n}{2}-\frac{x}{2}\sqrt{n}\right)}{n}\right]^{\frac{n}{2}-\frac{x}{2}\sqrt{n}}}\,.\,\frac{1}{\left[\dfrac{2\left(\frac{n}{2}+\frac{x}{2}\sqrt{n}\right)}{n}\right]^{\frac{n}{2}+\frac{x}{2}\sqrt{n}}}$$

Remplaçant $p(n-p)$ par sa valeur moyenne égale à $\frac{n^2}{2}$ il vient :

$$P(x)\,dx = \frac{1}{\sqrt{2\pi}}\,\frac{dx}{\left(1-\frac{x}{\sqrt{n}}\right)^{\frac{n}{2}-\frac{x}{2}\sqrt{n}}\,.\,\left(1+\frac{x}{\sqrt{n}}\right)^{\frac{n}{2}+\frac{x}{2}\sqrt{n}}}$$

$$= \frac{1}{\sqrt{2\pi}}\,\frac{dx}{\left(1-\frac{x^2}{n}\right)^{\frac{n}{2}}\left(1-\frac{x}{\sqrt{n}}\right)^{-\frac{x}{2}\sqrt{n}}\left(1+\frac{x}{\sqrt{n}}\right)^{\frac{x}{2}\sqrt{n}}}$$

Expression qui pour n infini tend vers :

$$\frac{1}{\sqrt{2\pi}}\,\frac{dx}{e^{-\frac{x^2}{2}}\times{}^{+}e^{\frac{x^2}{2}}\times e^{\frac{x^2}{2}}}$$

d'où finalement :

$$P(x)\,dx = \frac{1}{\sqrt{2\pi}}\,e^{-\frac{x^2}{2}}\,dx \qquad (1)$$

Ainsi lorsque l'on prend pour variable l'erreur réduite x, les erreurs accidentelles vraies tendent vers la distribution fameuse en e^{-x^2} connue sous le nom de loi de GAUSS ou de LAPLACE, sous la seule réserve que la composante moyenne ε soit très petite et que les signes + et — soient disposés de manière aléatoire. Il est à remarquer que si l'on peut par avance prédire la disposition de ces signes, c'est que l'erreur considérée est systématique par définition même.

Nous avons appelé la variable x l'erreur réduite et nous avons donné la définition plus haut. Nous allons en tirer une intéressante relation. Soit e une erreur accidentelle, par définition nous pouvons écrire :

$$e = \pm\,\varepsilon_1 \pm \varepsilon_2 \pm \dots \varepsilon_n.$$

Si le nombre des signes + est par exemple n-p, celui des signes — est p, d'où :

$$e \sim (n-p)\,\varepsilon_m - p\varepsilon_m = (n-2p)\,\varepsilon_m = x\varepsilon_m\sqrt{n}$$

ceci d'après la loi de formation de x.

L'erreur commise est donc bien proportionnelle à x, le facteur de proportionnalité étant $\varepsilon_m\sqrt{n}$ qui n'est autre de par sa formation même que la valeur moyenne quadratique de e.

Posons $x = h\nu\sqrt{2}$ d'où $dx = h\sqrt{2}\,.\,d\nu$, on trouve :

$$\mathrm{P}(\nu)\,d\nu = \frac{h}{\sqrt{\pi}}\,e^{-h^2\nu^2}\,d\nu$$

avec :

$$(2) \qquad \varepsilon_m\sqrt{n} = \frac{1}{h\sqrt{2}}$$

La quantité h est ce que l'on nomme le module de précision des mesures, on voit qu'il est lié à l'erreur moyenne quadratique d'une mesure isolée par la relation (2).

La probabilité pour qu'une erreur soit comprise entre ν et ν' est donc égale à :

$$P(\nu \,.\, \nu') = \frac{h}{\sqrt{\pi}} \int_{\nu}^{\nu'} e^{-h^2\nu^2} \, d\nu \tag{3}$$

Cette loi de probabilité est une loi à un seul paramètre h (ou $\varepsilon_m \sqrt{n}$) c'est la loi connue sous le nom de loi normale.

Introduction de la loi de Gauss par le postulat de la moyenne.

Le premier théorème de Bernouilli nous indique que la fréquence tend vers la probabilité et que par conséquent la moyenne arithmétique tend vers la valeur la plus probable de l'inconnue. Les écarts tendent donc vers les erreurs vraies.

L'étude de l'alternative répétée nous montre que la probabilité d'une erreur ν est de la forme

$$P(\nu)\, d\nu = \frac{h}{\sqrt{\pi}} e^{-h^2\nu^2} \, d\nu$$

On peut donc considérer comme établis mathématiquement les résultats annoncés au Chapitre II concernant les erreurs accidentelles et leur probabilité de répartition. On peut obtenir ces résultats d'une manière plus rapide en faisant appel au postulat des moyennes, supposant connue la loi des grands nombres.

Soit en effet $P(\nu)$ la probabilité de l'ensemble de tous les écarts $\nu_1\ \nu_2 \dots \nu_n$. Écrivons qu'elle est maxima :

$$P(\nu_1\ \nu_2 \dots \nu_n) = P(\nu_1) \,.\, P(\nu_2) \dots P(\nu_n)$$

et la condition de maximum nous impose que :

$$\frac{P'(\nu_1)}{P(\nu_1)} + \frac{P'(\nu_2)}{P(\nu_2)} + \dots + \frac{P'(\nu_n)}{P(\nu_n)} = 0. \text{ (dérivée logarithmique).}$$

Identifions cette expression avec celle du postulat de la moyenne :

$$C(\nu_1 + \nu_2 + \dots + \nu_n) = 0.$$

On est amené à poser :

$$\frac{P'(\nu)}{P(\nu)} = C\nu.$$

d'où en intégrant :

$$P(\nu) = Ke^{\frac{c\nu^2}{2}}.$$

La constante C est certainement négative puisque la probabilité d'une erreur infinie est nulle (inégalité de Bienaymé). On posera par conséquent $C = -h^2$. Il vient donc $P(\nu) = Ke^{-h^2\nu^2}$.

Dautre part K et h^2 ne sont pas indépendants. La probabilité pour qu'une erreur soit comprise entre $-\infty$ et $+\infty$ est une certitude, donc :

$$K\int_{-\infty}^{+\infty} e^{-h^2\nu^2} \, d\nu = 1.$$

Or on sait, c'est un résultat classique que $\int_{-\infty}^{+\infty} e^{-h^2\nu^2} d\nu = \frac{\sqrt{\pi}}{h}$, on a donc $K = \frac{h}{\sqrt{\pi}}$ et par conséquent :

$$P(\nu)\, d\nu = \frac{h}{\sqrt{\pi}} e^{-h^2\nu^2} \, d\nu$$

on retrouve bien là les résultats (1) et (3). Inversement d'ailleurs par identification, le calcul des probabilités nous permettrait de calculer la valeur de l'intégrale de Laplace si l'on ne connaissait pas d'autre méthode pour l'établir.

Calcul des erreurs moyennes.

Nous avons déjà vu que la moyenne quadratique d'une série d'erreurs $e_1, e_2, \ldots, e_n$ était égale à $\frac{1}{h\sqrt{2}}$. Nous allons retrouver ce résultat par un procédé inverse et lui en adjoindre un certain nombre d'autres.

Soit v la valeur d'une erreur : la quantité $v \, . \, \mathrm{P}(v)$ est égale (quelle que soit la loi de probabilité) à $v \times \frac{\mathrm{N}(v)}{\mathrm{N}}$ où $\mathrm{N}(v)$ est le nombre des erreurs égales à v et N le nombre total des erreurs.

La quantité $v \times \mathrm{N}(v)$ est donc la valeur de la somme des erreurs égales à v et l'erreur arithmétique est égale à :

$$\varepsilon_a = |\, v \,| \, . \, \mathrm{P}(v) = \frac{2h}{\sqrt{\pi}} \int_0^{+\infty} e^{-h^2v^2} . \, v dv.$$

Cette expression s'intègre immédiatement et donne :

$$\varepsilon_{m.a} = \frac{1}{h\sqrt{\pi}}.$$

On trouvera de même pour l'erreur moyenne quadratique :

$$\varepsilon_{mq}^2 = \int_{-\infty}^{+\infty} v^2 . \, \mathrm{P}(v) \, dv$$

pour la moyenne d'ordre 3 :

$$\varepsilon_3^3 = \int_{-\infty}^{+\infty} v^3 . \, \mathrm{P}(v) \, dv$$

pour la moyenne d'ordre 4 :

$$\varepsilon_4^4 = \int_{-\infty}^{+\infty} v^4 . \, \mathrm{P}(v) \, dv.$$

La méthode générale pour obtenir les valeurs numériques ainsi définies consiste à intégrer par parties et à établir des formules de récurrence entre les diverses intégrales, toutefois pour l'erreur moyenne quadratique, il est un artifice de calcul commode :

$$\varepsilon^2 = \frac{h}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-h^2v^2} . \, v^2 \, dv.$$

Considérons l'expression :

$$\frac{h}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-h^2v^2} \, dv = 1$$

et dérivons là par rapport au paramètre h. Il vient :

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-h^2v^2} \, dv + \frac{h}{\sqrt{\pi}} \int_{-\infty}^{+\infty} - 2hv^2 . \, e^{-h^2v^2} \, dv = 0$$

c'est-à-dire puisque :

$$\int_{-\infty}^{+\infty} e^{-h^2v^2} \, dv = \frac{\sqrt{\pi}}{h}$$

$$\frac{1}{h} - 2h \, . \, \varepsilon^2 = 0$$

(2) $$\varepsilon_{m.q} = \frac{1}{h\sqrt{2}}$$

c'est la valeur (2) déjà trouvée.

Les relations complètes sont les relations (4).

(4)
$$\Sigma \frac{|e|}{n} = \frac{1}{h\sqrt{\pi}} \qquad \Sigma \frac{e^2}{n} = \frac{1}{2h^2}$$
$$\Sigma \frac{|e^3|}{n} = \frac{1}{h^3\sqrt{\pi}} \qquad \Sigma \frac{e^4}{n} = \frac{3}{4h^4}$$

La vérification numérique de ces relations est un criterium qui permet de vérifier que les erreurs suivent la loi de GAUSS — en particulier on doit avoir pour un nombre assez grand d'erreurs :

(5)
$$\frac{\varepsilon^2_{m.q}}{\varepsilon^2_{m.a}} = \frac{\pi}{2}$$

Loi réduite. Tables de probabilités.

Nous avons déjà vu que l'on passait de l'expression $P(x, x') = \frac{1}{\sqrt{2\pi}} \int_x^{x'} e^{\frac{-x}{2}} dx$ à l'expression $P(v, v') = \frac{h}{\sqrt{\pi}} \int_v^{v'} e^{-h^2v^2} dv$ par un simple changement de variable.

Dans la pratique on fait plutôt l'inverse, on calcule l'erreur moyenne quadratique ε et l'on prend comme variable la quantité $\frac{v}{\varepsilon} \left(\text{où } \varepsilon = \frac{1}{h\sqrt{2}}\right)$. La variable $x = \frac{v}{\varepsilon}$ obéit en effet à la *loi réduite*.

(1)
$$P(x, x') = \frac{1}{\sqrt{2\pi}} \int_x^{x'} e^{\frac{-x^2}{2}} dx.$$

C'est ce que nous avons vérifié avec le 2e théorème de BERNOUILLI dans le cas du jeu de pile ou face.

La loi réduite peut être mise en table une fois pour toutes, table dans laquelle on rentre avec l'argument $\frac{v}{\varepsilon}$; en particulier, puisque la fréquence tend vers la probabilité, on peut voir immédiatement si une courbe des fréquences suit la loi de GAUSS.

Les tables de l'expression (1) sont les tables dites de la fonction $\Theta_0(x)$. Elles indiquent par un nombre décimal la fraction de l'aire totale de la courbe de GAUSS comprise entre l'abscisse zéro et l'abscisse $\frac{v}{\varepsilon} = x$ considérée.

Introduction de l'écart probable.

Beaucoup d'usagers, les artilleurs en particulier, utilisent comme unité de mesure « l'écart probable « dont nous avons vu la définition. C'est l'abscisse pour laquelle la demi-courbe de GAUSS est partagée en deux aires équivalentes.

Autrement dit :

$$\frac{1}{2} = \frac{2h}{\sqrt{\pi}} \int_0^{e_p} e^{-h^2v^2} dv$$
$$\frac{1}{2} = \frac{2}{\sqrt{\pi}} \left(he_p - \frac{1}{3} (he_p)^3 + \dots\right)$$

Cette équation donne par résolution, par approximations successives :

$$he_p = 0{,}48 \text{ d'où } e_p = 0{,}67\ \varepsilon = \frac{2}{3}\ \varepsilon. \tag{6}$$

L'écart probable est donc bien égal aux $\frac{2}{3}$ de l'erreur moyenne quadratique. On peut alors considérer la variable $\frac{\nu}{e_p}$ au lieu de $\frac{\nu}{\varepsilon}$ et considérer une courbe réduite où la variable soit :

$$x' = \frac{\nu}{e_p}.$$

La table de probabilité construite sur ce modèle s'utilise exactement comme la table de la fonction $\Theta_0(x)$. C'est celle qui est donnée au Chapitre II.

L'avantage de l'écart probable réside surtout dans sa liaison avec des valeurs numériques simples :

a) Il est plus facile à calculer que l'erreur moyenne quadratique ; par contre son domaine d'incertitude est plus grand.

b) Il est sensiblement égal au 1/4 de la dispersion (pour 100 événements), mais ce point n'est pas absolu et fixe seulement la zone où il faut le chercher.

c) pour 100 événements on peut admettre en gros que 50 sont « positifs », 50 « négatifs » et que la répartition interne des fréquences est la suivante (voir également Chapitre II) :

0 – e_p	e_p – $2e_p$	$2e_p$ – $3e_p$	$3e_p$ – $4e_p$
25	16	7	2

Ces nombres sont faciles à retenir pour les usagers, par contre pour ceux qui préfèrent des tables plus détaillées, l'usage de l'erreur moyenne quadratique est plus répandu. De toute façon les deux tables sont absolument équivalentes comme emploi.

Tolérances.

Si l'on consulte une table des probabilités pour les valeurs réduites, il est possible de définir une abscisse au-delà de laquelle l'écart a une probabilité très faible. Dans le cas des observations directes, on admet que le « seuil de signification » est dépassé lorsque la probabilité de cet écart est inférieure à 1 %. Si l'on consulte une table des écarts probables, on constate que tout écart $\nu \geqq 4e_p$ a une probabilité inférieure à 1 % (0,007), ce qui correspond sensiblement à $\nu \geqq 2{,}7\ \varepsilon$.

Si $\nu \geqq 3\varepsilon$ $P(\nu) = 0.003)$; on convient dans ces conditions, de considérer toute mesure discordante d'une quantité de cet ordre comme douteuse.

Il va de soi que ce criterium doit être élargi quand le nombre des mesures est très élevé.

Composition de deux lois normales.

Nous avons déjà vu que la composition de deux lois de probabilités quelconques fournissait une loi résultante dont l'erreur moyenne quadratique s'obtenait par la formule :

$$\eta^2(z) = \eta^2(x) + \eta^2(y).$$

Nous nous proposons de chercher la loi résultant de la composition de deux lois de Gauss et de montrer que cette loi obéit encore à une loi de Gauss.

Soient deux lois quelconques $f(x)$, $g(y)$ associées aux variables éventuelles x et y. Dans le domaine S la probabilité de $z = x + y$ sera de la forme $\iint_S f(x)\, g(y)\, dx\, dy$ en vertu du théorème des probabilités composées, étendu à tous les cas $z_p = x_i + y_j$ susceptibles de se produire. Pour effectuer la composition des deux lois proposées, il faut mettre cette expression sous la forme $K\int_{z'}^{z''} \varphi(z)\, dz$ où $\varphi(z)$ sera la loi de probabilité relative à $z = x + y$.

On devra donc effectuer un changement de variables tel que :

$$z = x + y$$
$$u = \lambda x + \mu y$$

λ et μ étant des coefficients numériques quelconques.

Dans ces conditions on aura :

$$\iint_S f(x)\, g(y)\, dx\, dy = \iint_{S'} h(z, u)\, dz\, du = \int_{z'}^{z''} dz \int_{-\infty}^{+\infty} h(z, u)\, du = \int_{z'}^{z''} \varphi(z)\, dz.$$

Si les lois f et g sont des lois normales on a :

$$\int f(x)\, g(y)\, dx\, dy = \int e^{-\left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right)}\, dx\, dy.$$

Prenons comme nouvelles variables $z = x + y$ et $u = \lambda x + \mu y$ en choisissant les coefficients λ et μ de manière que les directions $x + y$ et $\lambda x + \mu y$ soient conjuguées par rapport à l'ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2}$:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{A^2} + \frac{u^2}{B^2}.$$

Le déterminant fonctionnel qui intervient dans le changement de variables :

$$\frac{\partial(x, y)}{\partial(u, z)} = \frac{1}{\dfrac{\partial(u, z)}{\partial(x, y)}}$$

est toujours différent de zéro puisque :

$$\frac{\partial(u, z)}{\partial(x, y)} = \begin{vmatrix} 1 & 1 \\ \lambda & \mu \end{vmatrix}$$

Par conséquent :

$$\iint_S e^{-\left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right)}\, dx\, dy = \int_{z'}^{z''} dz \int_{-\infty}^{+\infty} e^{-\left(\frac{z^2}{A^2} + \frac{u^2}{B^2}\right)} \frac{du}{\begin{vmatrix} 1 & 1 \\ \lambda & \mu \end{vmatrix}} = C \int_{z'}^{z''} e^{-\frac{z^2}{A^2}}\, dz.$$

C'est bien là l'expression d'une loi normale, dans laquelle il serait facile de calculer C et A^2 et de s'assurer que l'erreur moyenne quadratique est en particulier la somme des carrés des erreurs composantes. Ce résultat porte le nom du théorème de d'Ocagne ; il était pourtant déjà connu de Cauchy, Poisson, mais avait été perdu de vue jusqu'à ce que d'Ocagne l'ait à nouveau démontré. Il montre qu'une somme d'erreurs accidentelles obéit à la loi normale, conclusion à laquelle nous étions déjà parvenus par une autre méthode.

Conclusion générale.

L'introduction du calcul des probabilités dans la théorie des erreurs accidentelles ouvre une voie féconde.

Il est possible grâce au premier théorème de BERNOUILLI de montrer la légimité de l'emploi de la moyenne arithmétique ; de préciser par le 2e théorème la loi normale de répartition des erreurs accidentelles et de passer à l'application numérique.

Des relations tout à fait nouvelles sont ainsi mises en évidence et permettent de juger si la loi normale est plus ou moins fidèlement suivie.

Si l'étude empirique des erreurs nous a montré l'existence de lois générales, seul le calcul des probabilités nous en a livré la forme, précisé la légimité et établi les bases sur un socle solide.