Some remarks on independent random variables.

By ALFRÉD RÉNYI in Budapest.

Let us consider a sequence of independent random variables $X_1, X_2, X_3, \ldots, X_n, \ldots$ having the same distribution function $F(x)$, with the mean value 0,

$$(1) \qquad \int_{-\infty}^{+\infty} x\,dF(x) = 0$$

and the finite dispersion σ,

$$(2) \qquad \int_{-\infty}^{+\infty} x^2\,dF(x) = \sigma^2.$$

Let us put

$$(3) \qquad S_n = X_1 + X_2 + \ldots + X_n$$

and let $F_n(x)$ denote the distribution function of S_n. It is well known[1]) that $F_{n+1}(x)$ is the convolution of $F_n(x)$ and $F(x)$, i. e.

$$(4) \qquad F_{n+1}(x) = \int_{-\infty}^{+\infty} F_n(x-v)\,dF(v).$$

In this case of "equal components" the central limit theorem is valid without any further supposition[2]), i. e. we have

$$(5) \qquad \lim_{n\to\infty} F_n(x\sigma\sqrt{n}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}.\,dt$$

for any fixed value of x. But the central limit theorem does not furnish any evidence regarding the asymptotic properties of $F_n(x\sigma\sqrt{n})$ if $|x|$ tends to ∞ together with n. Some information in this direction may be obtained from Liapounoff's theorem[3]) but only for the range of values $|X| = O(\sqrt{\log n})$. Considerably stronger results have been obtained by H. CRAMÉR[4]). One of his results[5]) is — under some additional conditions — valid for $|x| < c\sqrt{n}$, with some constant c. Evidently in general this is the maximal order of magnitude for which the problem has a sense, because if the random variables X_n are bounded, we have $S_n = O(n)$, and thus $F_n(x\sigma\sqrt{n}) = 1$ for $x > \lambda\sqrt{n}$ and $F_n(x\sigma\sqrt{n}) = 0$ for $x < -\lambda\sqrt{n}$, with some constant value of λ. In what follows we are concerned instead of the rather deep asymptotic laws with the more elementary question of estimating the value of constant λ, mentioned above, using only some particular data of $F(x)$.

Let ξ denote a bounded random variable, $F(x)$ its distribution function, let B denote the least number having the property that $F(x) = 1$ for $x > B$. Similarly let A denote the greatest number with the property that $F(x) = 0$ for $x < A$. Evidently B (resp. A) can also be defined as the least (resp. greatest) number for which the probability of $\xi > B$ (resp. of $\xi < A$) is equal to 0. We shall call B the probable least upper bound (*PLUB*) and A the probable greatest lower bound (*PGLB*) of ξ.

We start by proving the following

Lemma. Let $F(x)$ denote the distribution function of the bounded random variable ξ. Let us denote

$$m_k = \int_{-\infty}^{+\infty} x^k dF(x). \tag{6}$$

We suppose $m_1 = 0$, further we put

$$\Phi(x) = \int_{-\infty}^{x} F(t)\,dt. \tag{7}$$

If B and A denote the *PLUB* and *PGLB* of ξ, we have for any $k = 1, 2, \ldots$

$$B^{2k-1} - A^{2k-1} \geq \frac{m_{2k}}{\Phi(0)}. \tag{8}$$

Proof: It follows by partial integration, owing to $m_1 = 0$,

$$\Phi(B) = \int_A^B F(x)\,dx = [xF(x)]_A^B - \int_A^B x\,dF(x) = B. \tag{9}$$

Using (9) and applying partial integration twice, we obtain

$$m_{2k} = -(2k-1)B^{2k} + 2k(2k-1)\int_A^B x^{2k-2}\Phi(x)\,dx. \tag{10}$$

Owing to the fact, that the function $\Phi(x)$ is convex from below, we have

$$\begin{aligned} &\Phi(x) \leq -\frac{(x-A)\,\Phi(0)}{A} \quad \text{for} \quad A \leq x \leq 0 \\ &\Phi(x) \leq \Phi(0) + x\left(1 - \frac{\Phi(0)}{B}\right) \quad \text{for} \quad 0 \leq x \leq B. \end{aligned} \tag{11}$$

Applying the inequalities (11) we obtain from (10) the inequality (8). In what follows we shall use this Lemma only in the special case $k = 1$.

In this case we have evidently

$$B - A \geqq \frac{\sigma^2}{\Phi(0)}. \tag{12}$$

Owing to the convexity of $\Phi(x)$ we have $\Phi(0) \leqq \frac{-AB}{B-A}$ and thus it follows:

$$-AB \geqq \sigma^2. \tag{13}$$

As the arithmetic mean is greater than the geometric, we obtain the less precise inequality, (which has the advantage that it contains only σ),

$$B - A \geqq 2\sigma. \tag{14}$$

This is equivalent to a theorem proved first by J. L. WALSH[6]), an other proof has been given by A. HAAR[7]). Thus our lemma may be regarded as a sharpening of the result of WALSH. Now we prove the following

Theorem. Let $X_1, X_2, \ldots, X_n, \ldots$ denote independent random variables, all having the same distribution function $F(x)$, having the mean value 0 and dispersion σ. We put $S_n = X_1 + X_2 + \ldots + X_n$ and denote the *PLUB* of $|S_n|$ by M_n. Then it follows

$$M_n \geqq \frac{n\sigma^2}{2\Phi(0)}. \tag{15}$$

This theorem follows without any difficulty from (12), using the fact, that if B_n and A_n denote the *PLUB* and *PGLB* of S_n, we have by (4) $B_n = nB$ and $A_n = nA$, further we have $M_n = \operatorname{Max}(B_n, |A_n|)$. We have remarked above, that (14) is feebler than (12). Though actually we have deduced (14) from (12), this alone does not prove our assertion*), which has to be proved separately. As a matter of fact we have only to show, that

$$\sigma \geqq 2\Phi(0). \tag{16}$$

This can be proved as follows: We have by the inequality of SCHWARZ (for STIELTJES integrals)

$$\Phi(0) = \int_{-\infty}^{0} F(x)\,dx = -\int_{-\infty}^{0} x\,dF(x) \leqq \left(\int_{-\infty}^{0} x^2 dF(x) \,.\, \int_{-\infty}^{0} dF(x)\right)^{\frac{1}{2}} \tag{17}$$

Similarly, owing to $\int_{-\infty}^{+\infty} x\,dF(x) = 0$ we have

$$\Phi(0) = \int_{0}^{\infty} x\,dF(x) \leqq \left(\int_{0}^{\infty} x^2 dF(x) \,.\, \int_{0}^{\infty} dF(x)\right)^{\frac{1}{2}} \tag{18}$$

---

*) As a matter of fact, from $a \geqq b$ and $b \geqq \frac{c}{a}$ it follows $a \geqq \sqrt{c}$ but this does not imply $b \geqq \sqrt{c}$ (e. g. in the case $a = 3$, $b = 1$, $c = 2$).

Adding (17) and (18) and applying the inequality

$$(ab+cd)^2 \leq (a^2+c^2)(b^2+d^2)$$

we obtain (16).

Our theorem can be generalized for the case of unequal components, without any essential difficulty, the formulation of the corresponding theorem may be left to the reader. We mention that (15) is a "best possible" result, as there is equality in (15) for instance in the case of the binomial distribution.

An interesting application of our Theorem is the following: Let $1, \varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x), \ldots$ denote an orthonormal system of functions in the interval (0, 1). It is an often used evident fact, that

$$\sup_{0 \leq x \leq 1} |\varphi_i(x) - \varphi_j(x)| \geq \sqrt{2} \tag{19}$$

for any pair of indices $i \neq j$. Let us suppose in addition, that the functions $\varphi_n(x)$ are pairwise (stochastically) independent[8]), (as for instance the functions of the WALSH-system). It follows easily from (15) for $n=2$, combined with (16) that in this case we have

$$\sup_{0 \leq x \leq 1} |\varphi_i(x) - \varphi_j(x)| \geq 2. \tag{20}$$

This inequality is exact, as for instance there is equality in (20) for any pair of functions of the WALSH-system.

REFERENCES:

[1]) H. CRAMÉR, Mathematical Methods of Statistics, *Princeton, University Press*, 1946, p. 190.

[2]) ibidem p. 215.

[3]) H. CRAMÉR, Random variables and probability distributions, *Cambridge University Press*, 1937, p. 77.

[4]) H. CRAMÉR, Sur un nouveau théorème-limite de la théorie des probabilités, *Actualités Scientifiques et Industrielles*, No. 736, Paris, Hermann & Cie, 1938, pp. 5—23.

[5]) ibidem, Théorème 6, p. 20.

[6]) J. L. WALSH, A property of Haar's system of orthogonal functions, *Math. Annalen*, Vol. 90, 1923, p. 39.

[7]) A. HAAR, Über einige Eigenschaften der orthogonalen Funktionensysteme, *Math. Zeitschrift*, Vol. 31, 1930, p. 132.

[8]) For the definition see e. g. M. KAC, Sur les fonctions indépendantes I, *Studia Math.*, 6, 1936, pp. 46—58.