ON KOLMOGOROFF'S INEQUALITY

by

A. RÉNYI

§ 0. Notations

Let $S = [\Omega, \mathcal{A}, \mathbf{P}]$ be a probability space, i.e. Ω a set (the set of elementary events), $\mathcal{A}$ a σ-algebra of subsets of Ω, and $\mathbf{P}$ a probability measure on $\mathcal{A}$. We shall denote the elements of $\mathcal{A}$ (called random events) by capital letters and we denote by $\mathbf{P}(A)$ the probability of the event $A \in \mathcal{A}$. Random variables (i.e. functions defined on Ω and measurable with respect to $\mathcal{A}$) will be denoted by greek letters. We denote by $\mathbf{M}(\xi)$ the mean value and by $\mathbf{D}^2(\xi)$ the variance of the random variable ξ. We denote by $\mathbf{P}(A \mid B)$ the conditional probability of the event A with respect to the event B.

§ 1. Introduction

In the present paper we deal with the celebrated inequality of A. N. KOLMOGOROFF ([1]) according to which if $\xi_1, \xi_2, \ldots, \xi_n$ are independent random variables with mean value 0 and with finite variances $d_k^2 = \mathbf{D}^2(\xi_k)$ $(k = 1, 2, \ldots, n)$ then putting

$$\zeta_k = \xi_1 + \xi_2 + \ldots + \xi_k \qquad (k = 1, 2, \ldots n) \tag{1}$$

and

$$D_k^2 = d_1^2 + d_2^2 + \ldots + d_k^2 = \mathbf{D}^2(\zeta_k) \qquad (k = 1, 2, \ldots, n) \tag{2}$$

one has for any $\lambda > 1$

$$\mathbf{P}(\operatorname*{Max}_{1 \leq k \leq n} |\zeta_k| \geq \lambda D_n) \leq \frac{1}{\lambda^2} . \tag{3}$$

As well known, this inequality is extremely useful in proving the strong law of large numbers, the law of the iterated logarithm and other related theorems.

In § 2 we generalize this inequality by considering instead of (3) the conditional probability of the inequality $\operatorname*{Max}_{1 \leq k \leq n} |\zeta_k| \geq \lambda D_n$ with respect to some condition A having positive probability. We prove the following

Theorem. *If the random variables ξ_k are independent, have zero means, finite variances d_k^2 and finite fourth moments $f_k^4 = \mathbf{M}(\xi_k^4)$ $(k = 1, 2, \ldots, n)$, then if ζ_k resp. D_k are defined by* (1) *resp.* (2) *and we put*

$$F_n^4 = f_1^4 + \ldots + f_n^4 \tag{4}$$

then one has for any $\lambda > 1$, *and for any event* A *with* $\mathbf{P}(A)>0$,

$$\mathbf{P}(\operatorname*{Max}_{1\leq k\leq n}|\zeta_k| \geq \lambda D_n \mid A) \leq \frac{2+\sqrt{3+\left(\frac{F_n}{D_n}\right)^4}}{\lambda^2\sqrt{\mathbf{P}(A)}}. \tag{5}$$

It is known that in the proof of Kolmogoroff's inequality the supposition of independence of the random variables ξ_k can be replaced by the weaker supposition that the conditional mean value of ξ_k given $\xi_1, \ldots, \xi_{k-1}$ is identically equal to 0, that is that the variables ζ_k form a martingale (see [2]). It will be seen from the proof that the same supposition is sufficient for the validity of our Theorem.

§ 2. Proof of the generalization of Kolmogoroff's inequality

In this § we shall prove the Theorem formulated in § 1.

Let A be an arbitrary event, having positive probability $\mathbf{P}(A) > 0$. Let α denote the indicator of A, i.e. a random variable, which is equal to 1 on the set A (i.e. if the event A takes place) and equal to 0 on the complementary set $\bar{A} = \Omega - A$ (i.e. if the event A does not take place). Let B_k $(k=1, 2,\ldots, n)$ denote the event that $|\zeta_k|$ is the first term of the sequence $|\zeta_1|, |\zeta_2|, \ldots, |\zeta_n|$ which is not less than λD_n, i.e. B_k takes place if $|\zeta_1| < \lambda D_n, \ldots, |\zeta_{k-1}| < \lambda D_n$ and $|\zeta_k| \geq \lambda D_n$. Let β_k denote the indicator of B_k. Then clearly

$$0 \leq \sum_{k=1}^{n} \beta_k \leq 1, \text{ further } \beta_k \beta_l = 0 \text{ if } k < l \tag{6}$$

and β_k depends only on $\xi_1, \ldots, \xi_k$, and thus is independent of $\xi_{k+1}, \ldots, \xi_n$. Let finally C_n denote the event $\operatorname*{Max}_{1\leq k\leq n} |\zeta_k| \geq \lambda D_n$, that is C_n is the union of the sets $B_1, \ldots, B_n$. We have clearly

$$\begin{aligned}\mathbf{M}(\zeta_n^2\alpha) \geq \sum_{k=1}^{n}\mathbf{M}(\zeta_n^2\alpha\beta_k) = \sum_{k=1}^{n}\mathbf{M}(\zeta_k^2\beta_k\alpha) + 2\sum_{k=1}^{n}\mathbf{M}(\zeta_k\beta_k(\zeta_n-\zeta_k)\alpha) + \\ + \sum_{k=1}^{n}\mathbf{M}((\zeta_n-\zeta_k)^2\alpha\beta_k) \geq \sum_{k=1}^{n}\mathbf{M}(\zeta_k^2\beta_k\alpha) + 2\sum_{k=1}^{n-1}\sum_{j=k+1}^{n}\mathbf{M}(\zeta_k\beta_k\xi_j\alpha).\end{aligned} \tag{7}$$

Now put

$$\eta_{kj} = \zeta_k\beta_k\xi_j \qquad (1\leq k \leq n-1\,;\; k+1 \leq j \leq n). \tag{8}$$

Clearly we have, if $1 \leq k < j < h \leq n$

$$\mathbf{M}(\eta_{kj}\eta_{kh}) = \mathbf{M}(\zeta_k^2\beta_k\xi_j\xi_h) = \mathbf{M}(\zeta_k^2\beta_k)\,\mathbf{M}(\xi_j)\,\mathbf{M}(\xi_h) = 0, \tag{9a}$$

further if $k < l$, $k+1 \leq j$, $l+1 \leq h$ then owing to $\beta_k\beta_l = 0$ one has

$$\mathbf{M}(\eta_{kj}\eta_{lh}) = 0. \tag{9b}$$

Further

$$\mathbf{M}(\eta_{kj}^2) = \mathbf{M}(\zeta_k^2 \beta_k)\; d_j^2 . \tag{9c}$$

Thus the system

$$\eta_{kj}^* = \frac{\eta_{kj}}{d_j \sqrt{\mathbf{M}(\zeta_k^2 \beta_k)}} \tag{10}$$

is orthonormal. It follows by Bessel's inequality that

$$\left| \sum_{k=1}^{n} \sum_{j=k+1}^{n} \mathbf{M}(\eta_{kj}\, \alpha) \right| = \left| \sum_{k=1}^{n} \sum_{j=k+1}^{n} d_j \sqrt{\mathbf{M}(\zeta_k^2 \beta_k)} \cdot \mathbf{M}(\eta_{kj}^* \alpha) \right| \leqq \tag{11}$$

$$\leqq \sqrt{\mathbf{M}(\alpha^2)} \cdot \sqrt{\sum_{k=1}^{n-1} \mathbf{M}(\zeta_k^2 \beta_k) \sum_{j=k+1}^{n} d_j^2} \; .$$

Taking into account that

$$\mathbf{M}(\zeta_n^2 \beta_k) - \mathbf{M}(\zeta_k^2 \beta_k) = \mathbf{M}\big((\zeta_n - \zeta_k)^2 \beta_k\big) + 2\, \mathbf{M}\big(\zeta_k \beta_k (\zeta_n - \zeta_k)\big) \tag{12}$$

and $\mathbf{M}\big(\zeta_k \beta_k (\zeta_n - \zeta_k)\big) = 0$, it follows that

$$\mathbf{M}(\zeta_k^2 \beta_k) \leqq \mathbf{M}(\zeta_n^2 \beta_k). \tag{13}$$

Thus

$$\sum_{k=1}^{n-1} \mathbf{M}(\zeta_k^2 \beta_k) \left(\sum_{j=k+1}^{n} d_j^2 \right) \leqq D_n^2 \cdot \sum_{k=1}^{n-1} \mathbf{M}(\zeta_k^2 \beta_k) \leqq D_n^2\, \mathbf{M}(\zeta_n^2) = D_n^4. \tag{14}$$

Thus we obtain finally, taking into account that $\mathbf{M}(\alpha^2) = \mathbf{P}(A)$, that

$$\mathbf{M}(\zeta_n^2 \alpha) \geqq \sum_{k=1}^{n} \mathbf{M}(\zeta_k^2 \beta_k \alpha) - 2\, D_n^2 \sqrt{\mathbf{P}(A)}. \tag{15}$$

On the other hand if $\beta_k = 1$, one has $\zeta_k^2 \geqq \lambda^2 D_n^2$.
Thus

$$\sum_{k=1}^{n} \mathbf{M}(\zeta_k^2 \beta_k \alpha) \geqq \lambda^2 D_n^2\, \mathbf{M}\left(\alpha \left(\sum_{k=1}^{n} \beta_k \right) \right) = \lambda^2 D_n^2\, \mathbf{P}(A C_n) \tag{16}$$

where C_n stands for the event $\max\limits_{1 \leqq k \leqq n} |\zeta_k| \geqq \lambda D_n$. We obtain from (15) and (16)

$$\mathbf{P}(A C_n)\, \lambda^2 D_n^2 \leqq \mathbf{M}(\zeta_n^2 \alpha) + 2\, D_n^2 \sqrt{\mathbf{P}(A)}. \tag{17}$$

On the other hand,

$$\mathbf{M}(\zeta_n^2 \alpha) \leqq \sqrt{\mathbf{P}(A)\, \mathbf{M}(\zeta_n^4)} \, . \tag{18}$$

As clearly

$$\mathbf{M}(\zeta_n^4) \leqq F_n^4 + 3\, D_n^4 \tag{19}$$

we obtain from (17), (18) and (19)

$$\mathbf{P}(C_n \mid A) = \frac{\mathbf{P}(AC_n)}{\mathbf{P}(A)} \leq \frac{1}{\lambda^2 \sqrt{\mathbf{P}(A)}}\left(2 + \sqrt{3 + \frac{F_n^4}{D_n^4}}\right). \tag{20}$$

Thus (5) is proved.

Our theorem may e.g. be used to obtain an estimate for

$$\mathbf{P}(\operatorname*{Max}_{1 \leq k \leq \nu_n} |\zeta_k| \geq t_n)$$

where ν_n is a random variable, which may depend on the variables ξ_k. Let ν_n take on the values $n+1, n+2, \ldots, n+s$ with the corresponding probabilities $p_1, p_2, \ldots, p_s$. If A_l denotes the event $\nu_n = n + l$ $(l = 1, 2, \ldots, s)$ one has by Theorem 1, in the case $|\xi_k| \leq 1$ $(k = 1, 2, \ldots, n)$

$$\mathbf{P}(\operatorname*{Max}_{1 \leq k \leq \nu_n} |\zeta_k| \geq t_n) = \sum_{l=1}^{s} \mathbf{P}(\operatorname*{Max}_{1 \leq k \leq n+l} |\zeta_k| > t_n \mid A_l)\, \mathbf{P}(A_l) \leq \tag{21}$$

$$\leq \frac{4}{t_n^2} \sum_{l=1}^{n} \sqrt{\mathbf{P}(A_l)}\; D_{n+l}^2 \leq \frac{4\, D_{n+s}^2}{t_n^2} \sqrt{s}\,.$$

Thus we obtain, putting $t_n = \lambda\, D_{n+s}$ the following

Corollary. *If $\xi_1, \ldots, \xi_n$ are independent random variables, with mean value zero and satisfying $|\xi_k| \leq 1$, further if ν_n is a random variable capable of the values $n+1, \ldots, n+s$ and if D_k^2 denotes the variance of $\zeta_k = \xi_1 + \xi_2 + \ldots + \xi_k$, we have for $\lambda < 2\sqrt[4]{s}$*

$$\mathbf{P}(\operatorname*{Max}_{1 \leq k \leq \nu_n} |\zeta_k| > \lambda\, D_{n+s}) < \frac{4\sqrt{s}}{\lambda^2}. \tag{22}$$

(Received June 18, 1961.)

REFERENCES

[1] Kolmogoroff, A. N.: "Über die Summen durch den Zufall bestimmten unabhängigen Größen." *Math. Annalen* **99** (1929) 301—319.

[2] Doob, J. L.: *Stochastic processes.* Wiley, New-York, 1953.

О НЕРАВЕНСТВЕ А. Н. КОЛМОГОРОВА

A. RÉNYI

Резюме

Доказывается следующее обобщение известного неравенства А. Н. Колмогорова. Пусть ξ_k $(k = 1, 2, \ldots)$ независимые случайные величины, имеющие математическое ожидание 0, конечные дисперсии d_k и четвертые моменты f_k^4. Положим $\zeta_k = \xi_1 + \xi_2 + \ldots + \xi_k$, $D_n^2 = d_1^2 + d_2^2 + \ldots + d_n^2$,

$F_n^4 = f_1^4 + \ldots + f_n^4$. Пусть A произвольное событие с положительной вероятностью $\mathbf{P}(A) > 0$. Тогда имеет место для всех $\lambda > 1$

$$\mathbf{P}(\max_{1 \leq k \leq n} |\zeta_k| \geq \lambda D_n \mid A) \leq \frac{2 + \sqrt{3 + \left(\frac{F_n}{D_n}\right)^4}}{\lambda^2 \sqrt{\mathbf{P}(A)}}.$$