GENERALIZATION OF THE INEQUALITY OF MARKOFF

BY A. WALD

1. **Introduction.** Denote by X a random variable and by M_r the expected value $E \mid X - x_0 \mid^r$ of $\mid X - x_0 \mid^r$ for any integer r where x_0 denotes a given real value. M_r is also called the absolute moment of order r about the point x_0. For any positive number d, denote by $P(-d < X - x_0 < d)$ the probability that $\mid X - x_0 \mid < d$. The inequality of Markoff can be written as follows

$$(1) \qquad P(-d < X - x_0 < d) \geq 1 - \frac{M_r}{d^r}$$

The inequality (1) is also called, for $r = 2$, the inequality of Tchebyscheff. The inequality (1) can be written in the following way:

$$P(-\xi\sqrt[r]{M_r} < X - x_0 < \xi\sqrt[r]{M_r}) \geq 1 - \frac{1}{\xi^r}.$$

Substituting in the above inequality s for r and $\bar{\xi} \dfrac{\sqrt[r]{M_r}}{\sqrt[s]{M_s}}$ for ξ we get

$$(2) \qquad P(-\bar{\xi}\sqrt[r]{M_r} < X - x_0 < \bar{\xi}\sqrt[r]{M_r}) \geq 1 - \frac{1}{\bar{\xi}^s}\left(\frac{\sqrt[s]{M_s}}{\sqrt[r]{M_r}}\right)^s,$$

where r and s denote any integers and $\bar{\xi}$ denotes an arbitrary positive value.[1] Substituting in (2) $2k$ for s and 2 for r, we get the inequality of K. Pearson.[2] By other substitutions we get the formulae of Lurquin, Cantelli, etc.[3]

As is well known, the inequality (1) cannot be improved[4] for $d \geq \sqrt[r]{M_r}$. That is to say, to every $\epsilon > 0$ a random variable Y can be given such that

$$E \mid Y - x_0 \mid^r = E \mid X - x_0 \mid^r \quad \text{and} \quad P(-d < Y - x_0 < d) < 1 - \frac{M_r}{d^r} + \epsilon.$$

If the absolute moments $M_{i_1} = E(\mid X - x_0 \mid^{i_1}), \cdots, M_{i_j} = E \mid X - x_0 \mid^{i_j}$ of a random variable X are given (and no further data about X are known), then we shall say that a_d is the "*sharp*" *lower limit* of $P(-d < X - x_0 < d)$ if the following two conditions are fulfilled:

(1) For each random variable Y, for which $E \mid Y - x_0 \mid^{i_1} = E \mid X - x_0 \mid^{i_1}, \cdots, E \mid Y - x_0 \mid^{i_j} = E \mid X - x_0 \mid^{i_j}$, the inequality $P(-d < Y - x_0 < d) \geq a_d$ holds.

---

[1] The formula (2) has been given by A. Guldberg, *Comptes Rendus*, Paris, Vol. 175, p. 679.

[2] *Biometrika*, Vol. XII (1918–1919) pp. 284–296.

[3] E. Lurquin, *Comptes Rendus*, Paris, Vol. 175, p. 681. Also Cantelli, *Rendinconti delle Reale Academia dei Lincei*, 1916.

[4] See for instance, R. v. Mises, *Wahrscheinlichkeitsrechnung*, Leipzig, Vienna, Deuticke, 1931, p. 36.

(2) To each $\epsilon > 0$, a random variable Y can be given such that $E \mid Y - x_0 \mid^{i_\nu} = E \mid X - x_0 \mid^{i_\nu}$ ($\nu = 1, \cdots, j$) and $P(-d < Y - x_0 < d) < a_d + \epsilon$.

In other words, a_d is the *limes inferior*[5] of the probabilities $P(-d < Y - x_0 < d)$ formed for all random variables Y for which the i_ν-th absolute moment about the point x_0 is equal to the i_ν-th moment of X about the point x_0 ($\nu = 1, \cdots, j$).

PROBLEM: *The absolute moments $M_{i_1}, M_{i_2}, \cdots, M_{i_j}$ of a random variable X are given about the point x_0, where $i_1, i_2, \cdots, i_j$ denote any integers and M_{i_ν} denotes the moment of order i_ν ($\nu = 1 \cdots k$). It is required to calculate the "sharp" lower limit of the probability $P(-d < X - x_0 < d)$ for any positive value d.*

If only a single moment M_r is given, our problem is already solved, because the inequality (1) gives us the "sharp" lower limit for $d \geq \sqrt[r]{M_r}$ and for $d < \sqrt[r]{M_r}$ the "sharp" limit is obviously equal to zero. But the case in which even two moments M_r and M_s are given has not yet been solved. The formula (2) gives us a limit for $P(-d < X - x_0 < d)$, but this limit is not "sharp," as can easily be shown.

We shall give here some results concerning the general case, and the complete solution if only two moments M_r and M_s are given. We shall see that the "sharp" limit is considerably greater than the limit given by (2).

2. Some Propositions Concerning the General Case. We shall call a random variable X non-negative if $P(X < 0) = 0$. Since the absolute moments of the non-negative random variable $Y = \mid X - x_0 \mid$ about the origin are equal to the absolute moments of X about the point x_0 and since $P(Y < d) = P(-d < X - x_0 < d)$, the following proposition holds true:

(I) *Denote by $M_{i_1}, \cdots, M_{i_j}$ the absolute moments of order $i_1, \cdots, i_j$ of a certain random variable X about the point x_0. The limes inferior of the probabilities $P(-d < Y - x_0 < d)$ is equal to the limes inferior of the probabilities $P(Z < d)$, where $P(-d < Y - x_0 < d)$ is formed for all random variables Y for which the i_ν-th absolute moment about x_0 is equal to M_{i_ν} ($\nu = 1, \cdots, j$), and $P(Z < d)$ is formed for all non-negative random variables Z for which the i_ν-th moment about the origin is equal to M_{i_ν} ($\nu = 1, \cdots, j$).*

On account of the proposition (I) we can restrict ourselves to the consideration of non-negative random variables and of the moments about the origin.

A random variable X for which k different values $x_1, \cdots, x_k$ exist such that the probability $p(x_i)$ of x_i ($i = 1, \cdots, k$) is positive and $\sum_{i=1}^{k} p(x_i) = 1$, is called an *arithmetic* random variable of degree k. A random variable X will be called t-limited, if $P(-t \leq X \leq t) = 1$ We shall prove the following proposition.

(II). *Let us denote by $M_{i_1}, M_{i_2}, \cdots, M_{i_j}$ the absolute moments of order $i_1, \cdots, i_j$ of a certain non-negative random variable X, about the origin. Denote by $\Omega(k, t)$ the set of all non-negative t-limited arithmetic random variables of*

[5] The *limes inferior* of a set N of numbers is the greatest value y for which the inequality $y \leq x$ for each element x of N holds true. This is also called greatest lower bound.

degree $\leq k$, *for which the* i_ν*th moment about the origin is equal to* M_{i_ν} $(\nu = 1, \cdots, j)$. $\Omega(k, t)$ *is supposed to be not empty.*[6] *Denote further by* $a(d, k, t)$ *the limes inferior of the probabilities* $P(Y < d)$ *formed for all random variables* Y *of the set* $\Omega(k, t)$. *Then we can say: There exists in* $\Omega(k, t)$ *a random variable* Z *for which* $P(Z < d) = a(d, k, t)$. *If* $0 < a(d, k, t) < 1$ *and* Z *is a random variable in* $\Omega(k, t)$ *for which* $P(Z < d) = a(d, k, t)$, *then there exist at most* $j - 1$ *different positive values* $x_1, \cdots, x_{j-1}$ *such that* $x_i \neq d$, $x_i \neq t$ *and the probability* $p(x_i)$ *of* x_i *is positive* $(i = 1, 2, \cdots, j - 1)$.

At first we shall prove that there exists a random variable Z in $\Omega(k, t)$ such that $P(Z < d) = a(d, k, t)$. Since $a(d, k, t)$ is the *limes inferior* of $P(Y < d)$ formed for all random variables Y in $\Omega(k, t)$, there exists in $\Omega(k, t)$ a sequence $\{Z_i\}$ $(i = 1, 2, \cdots, \text{ad inf.})$ of random variables, such that $\lim_{i=\infty} P(Z_i < d) = a(d, k, t)$. Arranged in ascending order of magnitude, the values of Z_i which have a positive probability are denoted by $x_{i,1}, x_{i,2}, \cdots, x_{i,k_i}$. Since Z_i is a t-limited non-negative arithmetic random variable of degree $\leq k$, we have $k_i \leq k$ and $0 \leq x_{i,r} \leq t$ $(r = 1, \cdots, k_i)$. It follows easily from this fact that there exists a subsequence $\{Z_{i_\nu}\}$ $(\nu = 1, 2, \cdots, \text{ad inf.})$ of $\{Z_i\}$ with two properties: First, that the variables $Z_{i_1}, Z_{i_2}, \cdots$ are of the same degree (say s), that is to say $k_{i_\nu} = s$ $(\nu = 1, 2, \cdots, \text{ad inf.})$; and second, that the sequence $\{x_{i_\nu,r}\}$ $(\nu = 1, 2, \cdots, \text{ad inf.})$ converges for each integer $r \leq s$. Let us denote $\lim x_{i_\nu,r}$ by x_r $(r = 1, 2, \cdots, s)$, and the probability that Z_{i_ν} takes the value $x_{i_\nu,r}$ by $p_{i_\nu,r}$. It is obvious that there exists a subsequence $\{Z_{n_\nu}\}$ $(\nu = 1, 2, \cdots \text{ad inf.})$ of $\{Z_{i_\nu}\}$ such that the sequence $\{p_{n_\nu,r}\}$ converges with increasing ν. Let us denote $\lim_{\nu=\infty} p_{n_\nu,r}$ by p_r, $(r = 1, 2, \cdots s)$. Since $p_{i_\nu,1} + \cdots + p_{i_\nu,r} = 1$, $\sum_{r=1}^{s} p_r = 1$ must hold true. We consider the random variable Z for which the probability that $Z = x_r$ is equal to p_r $(r = 1, 2, \cdots s)$ and for which no values except $x_1, \cdots, x_r$ are possible. The random variable Z is obviously an element of $\Omega(k, t)$ and $P(Z < d) = a(d, k, t)$.

Let us consider the case in which $0 < a(d, k, t) < 1$ and denote by Z a random variable of $\Omega(k, t)$ for which $P(Z < d) = a(d, k, t)$. We shall prove that there exist at most $j - 1$ different positive values $x_1, \cdots, x_{j-1}$ such that $x_i \neq d$, $x_i \neq t$ and the probability $p(x_i)$ of x_i is positive $(i = 1, 2, \cdots, j - 1)$. In order to prove this statement we shall suppose that there exist j different positive points $x_1, \cdots, x_j$ such that $x_i \neq d$, $x_i \neq t$ and $p(x_i) > 0$ $(i = 1, 2, \cdots, j)$. Then we can write

$$\sum_{\nu=1}^{j} x_\nu^{i_1} p(x_\nu) = M_{i_1} - \sum x^{i_1} p(x)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\sum_{\nu=1}^{j} x_\nu^{i_j} p(x_\nu) = M_{i_j} - \sum x^{i_j} p(x),$$

[6] This is certainly the case, if we choose k and t great enough.

where the summations on the right hand sides are to be taken over all values of x which are different from $x_1, \cdots, x_j$ and for which $p(x) > 0$.

Since $P(Z < d) = a(d, k, t)$ and $0 < a(d, k, t) < 1$ by hypothesis, there exist two non-negative values b and c, such that $b < d$, $c \geq d$, $p(b) > 0$, and $p(c) > 0$.

We define a new arithmetic random variable Z' as follows: $p'(b) = p(b) - \epsilon$, $p'(c) = p(c) + \epsilon$, and for all other values $p'(x) = p(x)$, where $p'(x)$ denotes the probability that $Z' = x$, and ϵ a positive number less than $p(b)$. Z' is obviously a non-negative arithmetic variable of the same degree as Z. The moments about the origin of the order $i_1, i_2, \cdots, i_j$ of Z' will in general not be equal to the corresponding moments of Z. However this can be obtained by a small displacement of the points $x_1, \cdots, x_j$ into a system of neighboring points $\bar{x}, \cdots, \bar{x}_j$, provided that ϵ is small enough. In order to show this, we have only to prove that the functional determinant

$$\Delta = \begin{vmatrix} i_1\bar{x}_1^{i_1-1}, & \cdots, & i_1\bar{x}_j^{i_1-1} \\ i_2\bar{x}_1^{i_2-1}, & \cdots, & i_2\bar{x}_j^{i_2-1} \\ \cdots & \cdots & \cdots \\ i_j\bar{x}_1^{i_j-1}, & \cdots, & i_j\bar{x}_j^{i_j-1} \end{vmatrix} p'(x_1) \cdots p'(x_j)$$

of the functions $f_1(\bar{x}_1, \cdots, \bar{x}_j) = \sum_{\nu=1}^{j} \bar{x}_\nu^{i_1} p'(x_\nu), \cdots, f_j(\bar{x}_1, \cdots \bar{x}_j) = \sum_{\nu=1}^{j} \bar{x}_\nu^{i_1} p'(x_\nu)$ does not vanish at the point $\bar{x}_1 = x_1, \cdots, \bar{x}_j = x_j$. Since $p'(x_1), p'(x_2), \cdots p'(x_j)$ are not equal to zero, we have only to show that

$$\Delta^* = \begin{vmatrix} x_1^{i_1-1}, & \cdots, & x_j^{i_1-1} \\ \cdots & \cdots & \cdots \\ x_1^{i_j-1}, & \cdots, & x_j^{i_j-1} \end{vmatrix} = \begin{vmatrix} 1, & \cdots, & 1 \\ x_1^{i_2-i_1}, & \cdots, & x_j^{i_2-i_1}, \\ \cdots & \cdots & \cdots \\ x_1^{i_j-i_1}, & \cdots & x_j^{i_j-i_1} \end{vmatrix} x_1^{i_1-1} \cdots x_j^{i_1-1} \neq 0$$

where $i_2 - i_1, \cdots, i_j - i_1$ can be assumed positive by denoting by i_1 the smallest of the integers $i_1, i_2, \cdots, i_j$.

Let us consider the polynomial in x given by

$$R(x) = \begin{vmatrix} 1, & \cdots, & 1, & 1 \\ x_1^{i_2-i_1}, & \cdots & x_{j-1}^{i_2-i_1}, & x^{i_2-i_1} \\ \cdots & \cdots & \cdots & \cdots \\ x_1^{i_j-i_1}, & \cdots, & x_{j-1}^{i_j-i_1}, & x^{i_j-i_1} \end{vmatrix}$$

According to a well-known algebraic proposition, the number of positive roots of $R(x)$ is less than or equal to the number of changes of sign in the sequence of coefficients of $R(x)$. Since the number of changes of sign in $R(x)$ is obviously less than or equal to $j - 1$, the number of positive roots of $R(x)$ is also less than or equal to $j - 1$. On the other hand $x = x_1, \cdots, x = x_{j-1}$ are $j - 1$ positive roots of $R(x)$. Hence for any positive value $x \neq x_1, \neq x_2, \cdots,$

$\neq x_{j-1}$ the polynomial $R(x)$ does not vanish. Thus $R(x_j)$ and therefore also Δ^* and Δ are not equal to zero.

Let us denote by Z^* the random variable which we get from Z' by a small displacement of the points $x_1, \cdots, x_j$ into a system of neighboring points $\bar{x}_1, \cdots, \bar{x}_j$, such that the moment of order i_ν of Z^* about the origin becomes equal to M_{i_ν} $(\nu = 1, 2, \cdots, j)$. By choosing ϵ small enough we can obtain the values $\bar{x}_1, \cdots, \bar{x}_j$ as near to $x_1, \cdots, x_j$ as we like. In particular, ϵ can be chosen so small that $\bar{x}_1, \cdots, \bar{x}_j$ are positive numbers less than t, and $\bar{x}_i > d$ or $< d$ accordingly as $x_i >$ or $< d$. Then Z^* is obviously an element of $\Omega(k, t)$. But for Z^*

$$P(Z^* < d) = P(Z' < d) = P(Z < d) - \epsilon = a(d, k, t) - \epsilon$$

holds true, which is a contradiction because $a(d, k, t)$ is the *limes inferior* of $P(Y < d)$ formed for all random variables Y contained in $\Omega(k, t)$. Hence our assumption that there exist j different positive numbers $x_1, \cdots, x_j$, for which $x_i \neq d$, $x_i \neq t$ and $p(x_i) > 0$ $(i = 1, 2, \cdots, j)$, cannot be true, and the proposition II is proved in all its parts.

It follows from the proposition II that $a(d, k, t)$ is independent of k. On account of this fact and of the fact that any random variable X can be arbitrarily well approximated by arithmetic random variables, we get the proposition:

III. *Let us denote by $M_{i_1}, \cdots, M_{i_j}$ the moments about the origin of order $i_1, \cdots, i_j$ of a certain non-negative random variable. Denote by $\Omega(t)$ the set of all non-negative t-limited random variables, for which the i_ν-th moment about the origin is equal to M_{i_ν} $(\nu = 1, \cdots, j)$. Denote further by $a(d, t)$ the limes inferior of the probabilities $P(Y < d)$ formed for all random variables Y contained in $\Omega(t)$. Then we can say: There exists in $\Omega(t)$ a random variable Z for which $P(Z < d) = a(d, t)$. If $0 < a(d, t) < 1$ and Z is a random variable for which $P(Z < d) = a(d, t)$, then there exist at most $j - 1$ different positive numbers $x_1, \cdots, x_{j-1}$, such that $x_i \neq d$, $x_i \neq t$, and the probability that $Z = x_i$, is positive $(i = 1, 2, \cdots, j - 1)$*

It is obvious that $a(d, t)$ decreases monotonically with increasing t. Hence $\lim_{t=\infty} a(d, t)$ exists and it can be easily shown that:

$a(d, t)$ converges towards a_d if $t \to \infty$.

3. **Solution of the Problem if Only Two Moments are Given.** Let us denote by M_r and M_s the absolute moments respectively of order r and s about the point x_0 of a certain random variable X, where r and s $(r < s)$ denote any integers.

Let us first consider the case

$$(\alpha) \qquad \frac{M_r}{d^r} \leq \frac{M_s}{d^s}$$

It follows from (1) that

$$a_d \geq 1 - \frac{M_r}{d^r}$$

We shall show that $a_d = 1 - \frac{M_r}{d^r}$ if $\frac{M_r}{d^r} \leq 1$. For this purpose let us consider the arithmetic random variable Y_b of degree 3 defined as follows:

$$p(x_0 + d) = \frac{M_r}{d^r} - \frac{\epsilon}{2}, \qquad p(x_0 + d + b) = \frac{\epsilon}{2}\left(\frac{d}{d + b}\right)^r$$

$$p(x_0) = 1 - p(x_0 + d) - p(x_0 + d + b)$$

where ϵ is a positive number and $p(u)$ denotes the probability for $Y_b = u$. The r-th moment about x_0 of Y_b is obviously equal to M_r. On account of (α) the s-th moment of Y_b about x_0 is less than or equal to M_s for $b = 0$. On the other hand the s-th moment of Y_b about x_0 will be greater than M_s if b is sufficiently large. Hence there exists a non-negative value b_0 such that the s-th moment of Y_{b_0} is equal to M_s.

Since $P(-d < Y_{b_0} - x_0 < d) = 1 - \frac{M_r}{d^r} + \frac{\epsilon}{2} - \frac{\epsilon}{2}\left(\frac{d}{d + b_0}\right)^r < 1 - \frac{M_r}{d^r} + \frac{\epsilon}{2}$ and since ϵ can be chosen arbitrarily small, we have

$$a_d = 1 - \frac{M_r}{d^r}.$$

If $\frac{M_r}{d^r} \geq 1$, then a_d is equal to zero, because a_d decreases monotonically with decreasing d and $a_d = 0$ for $d = \sqrt[r]{M_r}$.

We have now to consider the case

$$(\beta) \qquad \frac{M_r}{d^r} > \frac{M_s}{d^s}$$

First we shall show that

$$(3) \qquad \frac{M_r}{d^r} < 1.$$

In fact, if $\frac{M_r}{d^r}$ were ≥ 1, then making use of (β) we have $\left(\frac{M_r}{d^r}\right)^{\frac{s}{r}} \geq \frac{M_r}{d^r} > \frac{M_s}{d^s}$, and hence $(M_r)^{\frac{s}{r}} > M_s$. But this is not possible, because according to the well-known inequalities between moments, $(M_r)^{\frac{s}{r}}$ is less than or equal to M_s. It follows from (3) and (β) that

$$(4) \qquad \frac{M_s}{d^s} < 1.$$

In order to calculate a_d, we shall apply the propositions found in section 2. On account of the proposition I, a_d is equal to the *limes inferior* of $P(Y < d)$

where $P(Y < d)$ is formed for all non-negative random variables Y for which the r-th moment about the origin is equal to M_r and the s-th moment about the origin is equal to M_s. Hence we can restrict ourselves to the consideration of non-negative random variables and of the moments about the origin.

We shall show that $0 < a(d, t)$ holds for any positive value t. In order to prove this, it is sufficient to show that $a_d > 0$ since $a(d, t) \geq a_d$. It follows from the inequality (1) that $a_d \geq 1 - \frac{M_r}{d^r}$. Since, according to (3), $\frac{M_r}{t^r} < 1$, we have $a_d > 0$, and therefore also

$$a(d, t) > 0 \tag{5}$$

Let us see whether $a(d, t) < 1$. If $M_s = (M_r)^{\frac{s}{r}}$, then, as is well-known, only a single non-negative random variable X exists for which the r-th moment about the origin is equal to M_r and the s-th moment is equal to $(M_r)^{\frac{s}{r}}$, namely the arithmetic random variable X of degree 1 for which the probability that $X = \sqrt[r]{M_r}$ is equal to 1. Since $\sqrt[r]{M_r} < d$, as can be seen from (3), we have $P(X < d) = 1$, and therefore $a_d = 1$. Hence in this case our problem is already solved and we have to consider only the alternative:

$$M_s = M_r^{\frac{s}{r}} + \sigma^2 \; (\sigma^2 > 0) \tag{6}$$

We shall show that $a(d, t) < 1$ for $t > \sqrt[r]{M_r} + d_r$. For this purpose let us consider the non-negative arithmetic random variable Y_ϵ of the degree 3 defined as follows:

$$p(\sqrt[r]{M_r}) = 1 - \epsilon, \; p(t) = \epsilon \frac{M_r}{t^r} < \epsilon \frac{M_r}{t^r} < \epsilon$$

$$p(0) = 1 - p(\sqrt[r]{M_r}) - p(t) = \epsilon - \epsilon \frac{M_r}{t^r},$$

where $p(u)$ denotes the probability for $Y_\epsilon = u$, and ϵ is a positive number < 1.

The r-th moment of Y_ϵ is equal to

$$M_r p(\sqrt[r]{M_r}) + t^r p(t) = M_r .$$

The s-th moment of Y_ϵ is given by the expression

$$A = M_r^{\frac{s}{r}} p(\sqrt[r]{M_r}) + t^s p(t) = (1 - \epsilon) M_r^{\frac{s}{r}} + \epsilon t^s \frac{M_r}{t^r}.$$

On account of (6), A is less than M_s for $\epsilon = 0$. For $\epsilon = 1$ we have

$$A = t^{s-r} M_r > d^{s-r} M_r .$$

Since from (β) $d^{s-r} M_r > M_s$, we have $A > M_s$ for $\epsilon = 1$. Hence there exists a positive value $\epsilon_0 < 1$ for which $A = M_s$. Thus the r-th moment of Y_{ϵ_0} is equal to M_r and the s-th moment of Y_{ϵ_0} is equal to M_s. We have

$$P(Y_{\epsilon_0} < d) = p(0) + p(\sqrt[r]{M_r}) = \epsilon - \epsilon \frac{M_r}{t^r} + 1 - \epsilon = 1 - \epsilon \frac{M_r}{t^r} < 1.$$

Hence

$$a(d, t) < 1. \tag{7}$$

On account of (5) and (7) it follows from proposition III, that there exists a non-negative arithmetic random variable X belonging to the set $\Omega(t)$ such that $P(X < d) = a(d, t)$ and there exists at most one positive value $\delta(\neq d, \neq t)$ with positive probability. Hence $a(d, t)$ is equal to the *limes inferior* of the probabilities $P(Y < d)$ formed for all non-negative arithmetic random variables Y which have the following two properties:

(1) The r-th moment about the origin is equal to M_r and the s-th moment about the origin is equal to M_s

(2) There exists at most a single positive value $\delta(\neq d, \neq t)$ with positive probability.

Denote by Z a non-negative t-limited random variable with the properties (1), (2), and for which $P(Z < d) = a(d, t)$. The following equations hold

$$\begin{aligned} p(0) + p(\delta) + p(d) + p(t) &= 1 \\ p(\delta)\delta^r + p(d)d^r + p(t)t^r &= M_r \\ p(\delta)\delta^s + p(d)d^s + p(t)t^s &= M_s \end{aligned} \tag{8}$$

where $p(u)$ denotes the probability that $Z = u$.

From the last two equations of (8), we get

$$p(\delta) = \frac{M_r d^{s-r} - M_s + p(t)\,[t^s - t^r d^{s-r}]}{\delta^r(d^{s-r} - \delta^{s-r})} \tag{9}$$

$$p(d) = \frac{M_s - \delta^{s-r} M_r + p(t)\,[t^r \delta^{s-r} - t^s]}{d^r(d^{s-r} - \delta^{s-r})}. \tag{10}$$

Since $\frac{M_r}{d^r} > \frac{M_s}{d^s}$ and $t > d$, the numerator in (9) is positive. Since $0 \leq p(\delta) \leq 1$, the inequality

$$0 < \delta < d \tag{11}$$

must hold. Hence

$$p(\delta) > 0. \tag{12}$$

We shall show that $p(t) = 0$ if t is sufficiently large. For this purpose let us make the assumption $p(t) > 0$. We define a new random variable Z' as follows:

$$p'(t) = p(t) - \epsilon \text{ where } 0 < \epsilon < p(t)$$

$$p'(d) = p(d) - \epsilon \frac{t^r\delta^{s-r} - t^s}{d^r(d^{s-r} - \delta^{s-r})}$$

$$p'(\delta) = p(\delta) - \frac{\epsilon(t^s - t^r d^{s-r})}{\delta^r(d^{s-r} - \delta^{s-r})}$$

$$p'(0) = 1 - p'(\delta) - p'(d)\_p'(t)$$

and

$$p'(z) = 0 \text{ for all values } z \neq 0, \neq\delta, \neq d, \neq t.$$

$p'(u)$ denotes the probability that $Z' = u$.

The equations (8) remain satisfied if we substitute $p'(0)$, $p'(\delta)$, $p'(d)$, and $p'(t)$ for $p(0)$, $p(\delta)$, $p(d)$, and $p(t)$ respectively. Hence the r-th moment of Z' is equal to M_r and the s-th moment is equal to M_s. We have to show that Z' is in fact a random variable, that is to say, that the defined probabilities are ≥ 0 and ≤ 1. It is sufficient to show that the defined probabilities are non-negative, because the sum of them is equal to 1 and therefore they must be ≤ 1.

Obviously $p'(t)$ is >0. Since $t > d$ and according to (11) $d > \delta$, we have $p'(d) > p(d) > 0$. According to (12), $p(\delta)$ is positive. Hence for ϵ sufficiently small $p'(\delta)$ is also positive. We have to show that also $p'(0) \geq 0$. $p'(0)$ is given by

$$\begin{aligned} p'(0) &= 1 - p'(\delta) - p'(d) - p'(t) \\ &= 1 - p(\delta) - p(d) - p(t) + \epsilon\left[1 + \frac{t^r\delta^{s-r} - t^s}{d^r(d^{s-r} - \delta^{s-r})} + \frac{t^s - t^r d^{s-r}}{\delta^r(d^{s-r} - \delta^{s-r})}\right] \\ &= p(0) + \epsilon\,\frac{d^r\delta^r(d^{s-r} - \delta^{s-r}) + t^s(d^r - \delta^r) - t^r(d^s - \delta^s)}{d^r\delta^r(d^{s-r} - \delta^{s-r})}. \end{aligned}$$

Since $p(0) \geq 0$, $\epsilon > 0$, $d > \delta$ and $s > r$, this last expression is positive if t is sufficiently large. We may assume t so great that $p'(0) \geq 0$, because we want to calculate only

$$a_d = \lim_{t=\infty} a(d, t).$$

Now we shall show that

$$p'(d) + p'(t) > p(d) + p(t).$$

In fact

$$\begin{aligned} p'(d) + p'(t) - p(d) - p(t) &= \epsilon\left[\frac{t^s - t^r\delta^{s-r}}{d^r(d^{s-r} - \delta^{s-r})} - 1\right] \\ &= \epsilon\left[\frac{t^r}{d^r}\,\frac{t^{s-r} - \delta^{s-r}}{d^{s-r} - \delta^{s-r}} - 1\right] > 0. \end{aligned}$$

Then

$$p'(0) + p'(\delta) < p(0) + p(\delta) = a(d, t)$$

must follow. Since $p'(0) + p'(\delta) = P(Z' < d)$, we have a contradiction and therefore the assumption $p(t) > 0$ is reduced to an absurdity. Hence $p(t)$ must be equal to zero and $a(d, t) = a_d$. If we substitute zero for $p(t)$ in (8), (9), and (10) we obtain:

$$\begin{cases} p(0) + p(\delta) + p(d) = 1 \\ p(\delta)\delta^r + p(d)d^r = M_r \\ p(\delta)\delta^s + p(d)d^s = M_s \end{cases} \tag{13}$$

$$p(\delta) = \frac{M_r d^{s-r} - M_s}{\delta^r(d^{s-r} - \delta^{s-r})} \tag{14}$$

$$p(d) = \frac{M_s - M_r\delta^{s-r}}{d^r(d^{s-r} - \delta^{s-r})}. \tag{15}$$

We shall prove that $p(0) = 0$. For this purpose let us make the assumption $p(0) > 0$. Denote by δ_1 a positive number $<\delta$ and let us consider the arithmetic random variable Z' of degree 3 defined as follows:

$$p'(\delta_1) = \frac{M_r d^{s-r} - M_s}{\delta_1^r(d^{s-r} - \delta_1^{s-r})}$$

$$p'(d) = \frac{M_s - M_r\delta_1^{s-r}}{d^r(d^{s-r} - \delta_1^{s-r})}$$

$$p'(0) = 1 - p'(\delta_1) - p'(d).$$

The r-th moment of Z' is evidently equal to M_r and the s-th moment to M_s. Since $p(\delta) > 0$ according to (12), and $p(0) > 0$ by hypothesis, $p'(0)$ and $p'(\delta_1)$ will be greater than zero if δ_1 is sufficiently near to δ. The derivative of $p'(d)$ with respect to δ_1 is given by

$$\frac{1}{d^r}\,\frac{-M_r(s-r)\delta_1^{s-r-1}(d^{s-r} - \delta_1^{s-r}) + (s-r)\delta_1^{s-r-1}(M_s - M_r\delta_1^{s-r})}{(d^{s-r} - \delta_1^{s-r})^2}$$

$$= \frac{(s-r)\delta_1^{s-r-1}}{d^r(d^{s-r} - \delta_1^{s-r})^2}\,(M_s - M_r d^{s-r}).$$

Since $\frac{M_r}{d^r} > \frac{M_s}{d^s}$, the above expression is negative. Hence $p'(d)$ decreases with increasing δ_1. Since $\delta_1 < \delta$, we have

$$p'(d) > p(d) \geq 0$$

and therefore

$$1 - p'(d) < 1 - p(d) = a_d .$$

Since $1 - p'(d) = P(Z' < d)$, we have a contradiction and the assumption $p(0) > 0$ is proved an absurdity. Hence $p(0) = 0$, and $p(\delta) + p(d) = 1$. From (13), (14) and (15) we have

$$q(\delta) + p(d) = \frac{M_r d^s - M_s d^r + M_s \delta^r - M_r \delta^s}{\delta^r d^r (d^{s-r} - \delta^{s-r})} = 1 .$$

Hence

$$M_r d^s - M_s d^r + \delta^r (M_s - d^s) + \delta^s (d^r - M_r) = 0. \tag{16}$$

The equation (16) in δ has at most two positive roots, because the derivative of the left hand side of (16)

$$r\delta^{r-1}(M_s - d^s) + s\delta^{s-1}(d^r - M_r)$$

has exactly one positive root in δ. Since $\delta = d$ is a root of (16), the value of δ which we are seeking must be the second positive root of (16), which we shall denote by δ_0.

It can be easily shown that $\delta_0 \leq \sqrt[r]{M_r} < d$. In fact, for $\delta = 0$ the left hand side of (16) is positive on account of the assumption (β) and for $\delta = \sqrt[r]{M_r}$ it becomes equal to

$$M_s(M_r - d^r) - M_r^{\frac{s}{r}}(M_r - d^r) = \left(M_s - M_r^{\frac{s}{r}}\right)(M_r - d^r)$$

Since $M_s \geq M_r^{\frac{s}{r}}$ and recalling from (3) that $M_r < d^r$, the above expression is less than or equal to 0. Hence δ_0 lies between 0 and $\sqrt[r]{M_r} < d$.

Hence a_d is given by the expression

$$a_d = 1 - p(d) = 1 - \frac{M_s - M_r \delta_0^{s-r}}{d^r(d^{s-r} - \delta_0^{s-r})} . \tag{17}$$

For $s = 2r$ the root δ_0 can be easily calculated. We get

$$\delta_0 = \sqrt[r]{\frac{M_{2r} - d^r M_r}{M_r - d^r}} \tag{18}$$

If we substitute in (17) $2r$ for s and the right hand side of (18) for δ_0, then we get

$$\begin{aligned} a_d &= 1 - \frac{M_{2r} - M_r\left(\dfrac{M_{2r} - d^r M_r}{M_r - d^r}\right)}{d^r\left(d^r - \dfrac{M_{2r} - d^r M_r}{M_r - d^r}\right)} \\ &= 1 - \frac{(M_r - d^r)M_{2r} - M_r(M_{2r} - d^r M_r)}{d^r[d^r(M_r - d^r) - M_{2r} + M_r d^r]} \\ &= 1 - \frac{d^r(M_r^2 - M_{2r})}{d^r[2M_r d^r - d^{2r} - M_{2r}]} \\ &= 1 - \frac{M_r^2 - M_{2r}}{2M_r d^r - d^{2r} - M_{2r}} . \end{aligned}$$

Let us denote the non-negative number $M_{2r} - M_r^2$ by σ^2, then we obtain[7]

$$a_d = 1 - \frac{\sigma^2}{(d^r - M_r)^2 + \sigma^2}. \qquad (\sigma^2 = M_{2r} - M_r^2). \tag{19}$$

Let us compare the "sharp" limit given by (19) with the limit given by (2). If we substitute, in (2), $2r$ for s and d for $\xi\sqrt[r]{M_r}$ we have

$$b_d = 1 - \frac{M_{2r}}{d^{2r}} = 1 - \left(\frac{M_r}{d^r}\right)^2 - \frac{\sigma^2}{d^{2r}}$$

as a lower limit of the probability $P(-d < X < x_0 < d)$. We see that for small values of σ^2, b_d is considerably smaller than a_d.

Our results may be summarized in the following

THEOREM: *Denote by M_r the r-th and by M_s the s-th absolute moment of a random variable X about the point x_0, where $r < s$. For any positive value d denote by $P(-d < X < x_0 < d)$ the probability that $|X - x_0| < d$. The "sharp" lower limit a_d of $P(-d < X - x_0 < d)$ is defined as the limes inferior of the probabilities $P(-d < Y - x_0 < d)$ formed for all random variables Y for which the r-th moment about x_0 is equal to M_r and the s-th moment about x_0 is equal to M_s. We have to distinguish two cases.*

I. $\frac{M_r}{d^r} \leq \frac{M_s}{d^s}$. *In this case* $a_d = 1 - \frac{M_r}{d^r}$ *if* $\frac{M_r}{d^r} \leq 1$, *and* $a_d = 0$ *if* $\frac{M_r}{d^r} > 1$.

II. $\frac{M_r}{d^r} > \frac{M_s}{d^s}$. In this case a_d is given by

$$a_d = 1 - \frac{M_s - M_r\delta_0^{s-r}}{d^r(d^{s-r} - \delta_0^{s-r})}, \tag{17}$$

where δ_0 is the positive root $\neq d$ of the equation[8] in δ.

$$M_r d^s - M_s d^r + \delta^r(M_s - d^s) + \delta^s(d^r - M_r) = 0.$$

For $s = 2r$ we have

$$\delta_0 = \sqrt[r]{\frac{M_{2r} - d^r M_r}{M_r - d^r}}.$$

If we substitute in (17) $2r$ for s and the above expression for δ_0, we obtain

$$a_d = 1 - \frac{\sigma^2}{(d^r - M_r)^2 + \sigma^2},$$

where $\sigma^2 = M_{2r} - M_r^2$.

COWLES COMMISSION, COLORADO SPRINGS.

[7] The case $s = 2r$ has been treated also by Cantelli. He demonstrated the formula (19) in quite another way, which cannot be generalized for the case $s \neq 2r$. Cantelli's formula and its demonstration are given in the book of M. Fréchet, *Generalities sur Probabilities. Variables Aleatoires*, Paris, 1937, pp. 123–126.

[8] As has been shown, there exists exactly one positive root $\neq d$ of the equation considered.