ARKIV FÖR MATEMATIK, ASTRONOMI OCH FYSIK.
BAND 29 A. N:o 20.

Determination of the maximum deviation from the Gaussian law.

By

CARL-GUSTAV ESSEEN.

With 3 figures in the text.

Communicated February 10th 1943 by TORSTEN CARLEMAN and ARNE BEURLING.

Let $x_1, x_2, \ldots x_n$ be a sequence of independent random variables with the same distribution function $F(x)$, the mean value zero and the finite dispersion $\sigma \neq 0$. We denote by α_k the moment, by β_k the absolute moment of order k, $(k=1,2,3,\ldots)$, by $F_n(x)$ the distribution function corresponding to the variable $\frac{x_1+x_2+\ldots+x_n}{\sigma\sqrt{n}}$ and by $f(t)$ and $f_n(t)$ the characteristic function of $F(x)$ and $F_n(x)$ respectively. It is well-known that $f_n(t)=\left\{f\left(\frac{t}{\sigma\sqrt{n}}\right)\right\}^n$.

As $n \to \infty$, $F_n(x)$ converges to the normal distribution function $\Phi(x)$. Let us for a moment suppose that β_k is finite for every k. Subject to the condition that $\overline{\lim_{t\to\pm\infty}}|f(t)|<1$ CRAMÉR[1] has proved the following expansion:

$$(1) \qquad F_n(x)=\Phi(x)+\sum_{\nu=1}^{k-3}\frac{p_\nu(x)}{n^{\nu/2}}e^{-\frac{x^2}{2}}+O\left(\frac{1}{n^{\frac{k-2}{2}}}\right), \quad (n\to\infty),$$

k being an arbitrary integer ≥ 3 and $p_\nu(x)$ certain polynomials,

[1] H. CRAMÉR, On the composition of elementary errors. Skand. Aktuarietidskr. 11 (1928), p. 59. It is shown here that (1) is valid if $\beta_k<\infty$ for some $k \geqq 3$.

the coefficients of which are determined by the moments. If, however, $\overline{\lim\limits_{t\to\pm\infty}}|f(t)| = 1$, this being the case e. g. when $F(x)$ is purely discontinuous, things are changed. In a previous paper[2] I have shown that under all circumstances

$$(2) \qquad |F_n(x) - \Phi(x)| \leq C \cdot \frac{\beta_3}{\sigma^3 \sqrt{n}},$$

C being an absolute constant, but that an expansion such as (1) generally does not hold for $k > 3$.

Let us say that $F(x)$ belongs to (A) if $F(x)$ is a purely discontinuous function and the distance between any two discontinuities is an entire multiple of a constant positive quantity. In the paper cited above[3] I also proved that if $F(x) \not< (A)$ and β_3 is finite then

$$(3) \qquad F_n(x) = \Phi(x) + \frac{\alpha_3}{6\sigma^3\sqrt{2\pi n}}(1-x^2)e^{-\frac{x^2}{2}} + o\left(\frac{1}{\sqrt{n}}\right), \quad (n\to\infty),$$

If, however, $F(x) < (A)$ then $F_n(x)$ has in the vicinity of $x = 0$ discontinuities where the saltus is of the same order of magnitude as $\frac{1}{\sqrt{n}}$. Thus, if $F(x) < (A)$, there must enter into the expansion of $F_n(x)$ a discontinuous function with saltus of order of magnitude $\frac{1}{\sqrt{n}}$. It is just this function which renders an expansion such as (1) impossible in the general case, even if all absolute moments are finite.

In this note I shall state the expansion of $F_n(x)$ obtained when $F(x) < (A)$, and sketch a proof of it. As the expansion (3) holds for $F(x) \not< (A)$, we thus completely know the asymptotic behaviour of $F_n(x)$ up to and including the term of order of magnitude $\frac{1}{\sqrt{n}}$. Then we are able to determine the asymptotic maximum deviation of $F_n(x)$ from $\Phi(x)$. If $F(x)$ is symmetric and continuous for $x = 0$ we prove that

$$(4) \qquad \lim_{n\to\infty} \operatorname*{Max}_{-\infty<x<\infty} \sqrt{n}\,|F_n(x) - \Phi(x)| \leq \frac{1}{\sqrt{2\pi}}.$$

---

[2] C. G. Esseen, On the Liapounoff limit of error in the theory o probability. Arkiv för Mat. Astr. o. Fys. 28 A (1942), No. 9, p. 12.

[3] Cf. loc. cit.[2], theorem 3.

There is equality if and only if $F(x)$ is the symmetric BERNOULLI distribution.

1. For later purposes it is of importance to find out whether there exists a finite value $t = t_0 \neq 0$ such that $|f(t_0)| = 1$. This is answered to by the following lemma.[4]

Lemma. *If and only if* $F(x) < (A)$ *there exists a finite* $t_0 \neq 0$ *such that* $|f(t_0)| = 1$.

The condition is necessary, for suppose that $t_0 \neq 0$ is finite and $|f(t_0)| = f(0) = 1$. Then $f(t_0)\, e^{i\theta_0} = f(0)$ for some real θ_0 or $\int\limits_{-\infty}^{\infty} (1 - e^{i(\theta_0 + t_0 x)})\, d\, F(x) = 0$. On taking the real part we obtain $\int\limits_{-\infty}^{\infty} g(x)\, d\, F(x) = 0$, where $g(x) = 1 - \cos(\theta_0 + t_0 x)$. As $g(x) \geq 0$ and continuous, $g(x)$ must be zero at every point x where $d\, F(x) > 0$. But $g(x) = 0$ only for

$$(5) \qquad x = x_0 + \nu \cdot \frac{2\pi}{t_0}, \quad \left(x_0 = -\frac{\theta_0}{t_0}, \;\; \nu = 0, \pm 1, \pm 2, \ldots\right),$$

and thus $F(x)$ must be a purely discontinuous function with the saltus $a_\nu \geq 0$ only for

$$(5') \qquad x = x_0 + \nu \cdot \frac{2\pi}{t_0}.$$

Thus $F(x) < (A)$. The condition is sufficient, for if $F(x) < (A)$ let it be defined by $(5')$. Then $f(t) = \sum\limits_{\nu} a_\nu\, e^{it\left(x_0 + \nu \cdot \frac{2\pi}{t_0}\right)}$ and then $|f(t)|$ is periodic with the period t_0. Hence $|f(t_0)| = f(0) = 1$.

If $F(x) < (A)$ and t_0 is the least positive number such that $F(x)$ may be defined by $(5')$, we say that $F(x)$ belongs to $(A;\, t_0)$. Then $|f(t)|$ is periodic with the period t_0.[5]

2. In what follows we assume that the mean value of $F(x)$ is zero, that $\sigma \neq 0$ and β_3 finite but nothing is assumed about the higher absolute moments. In this section we also

[4] A. WINTNER, On a class of Fourier transforms. American Journ. of Math. 58 (1936), p. 49.

[5] It may happen that $|f(t)| = 1$ for every t. Then it is easily shown that $F(x) \equiv E(x - a) = \begin{cases} 0 & \text{for } a > x \\ 1 & a \leq x \end{cases}$. Since, however, we always suppose that $\alpha_1 = 0$ and $\sigma \neq 0$, this case is excluded here.

suppose that $F(x) < (A)$, and we shall determine the discontinuous function mentioned in the introduction.

In order to exemplify how this discontinuous function is formed let us consider the following case, the symmetric BERNOULLI distribution, $F(x)$ having the jump $\frac{1}{2}$ for $x = \pm \frac{1}{2}$. Here $F(x) < (A; 2\pi)$ and $f(t) = \cos t/2$. Suppose that n is even. Then $F_n(x)$ is a purely discontinuous function with discontinuity points $x = \frac{k}{\sqrt{n}}$, $(k = 0, \pm 2, \pm 4, \ldots, \pm n)$. As is well-

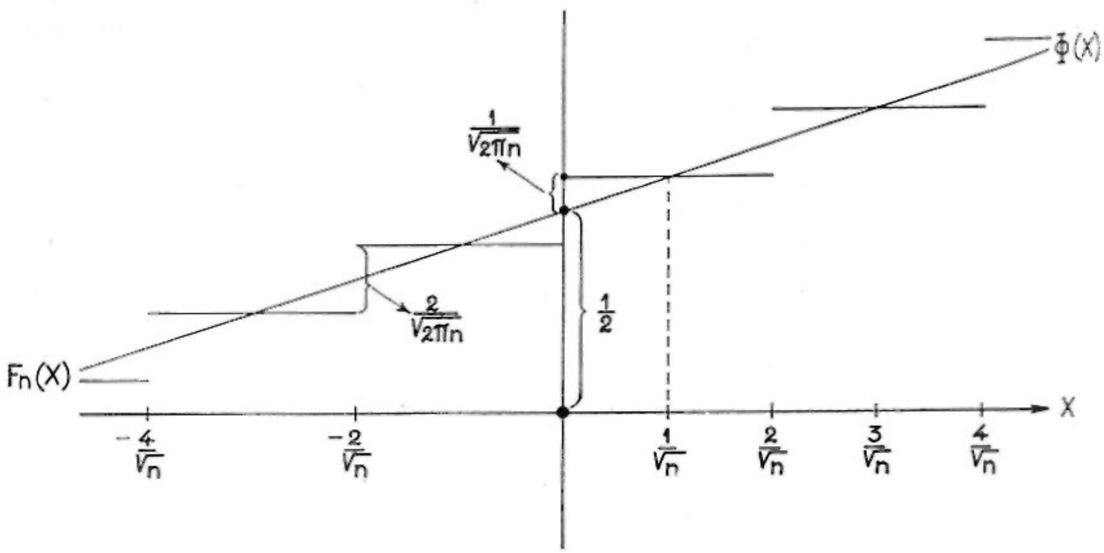

Fig. 1.

known, $F_n(x)$ has for a bounded discontinuity point x a saltus asymptotically equal to

$$(6) \qquad \frac{2}{\sqrt{2\pi n}} e^{-\frac{x^2}{2}}, \quad (n \to \infty).$$

Thus in the vicinity of $x = 0$ the jump of $F_n(x)$ and the growth of $\Phi(x)$ on an interval of length $\frac{2}{\sqrt{n}}$ are equal to $\frac{2}{\sqrt{2\pi n}} + o\left(\frac{1}{\sqrt{n}}\right)$. Further $\frac{1}{2}(F_n(+0) + F_n(-0)) = \frac{1}{2}$. Hence the asymptotic behaviour of $F_n(x)$ and $\Phi(x)$ may be represented by fig. 1 and 2. $\left(\text{Here we have neglected the term } o\left(\frac{1}{\sqrt{n}}\right)\right)$.

In fig. 1 and 2 it is easily seen that

(7) $$F_n(x) - \Phi(x) \sim \frac{2}{\sqrt{2\pi n}} P_1\left(\frac{x\sqrt{n}}{2}\right)$$

for small values of x, where

(8) $$P_1(x) = [x] - x + \frac{1}{2}$$

and $[x]$ is the integral part of x. On account of (6) and (7) let us form the function

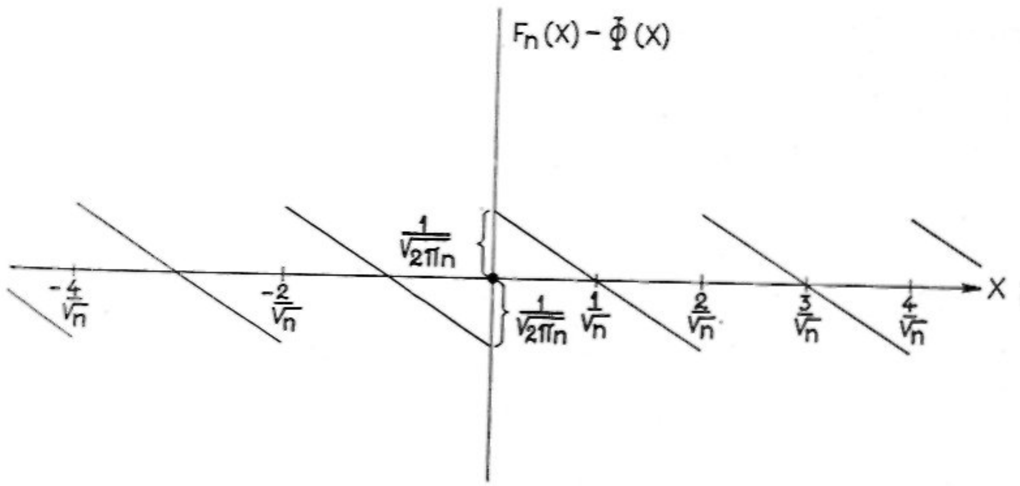

Fig. 2.

(9) $$D_n(x) \equiv \frac{2}{\sqrt{2\pi n}} P_1\left(\frac{x\sqrt{n}}{2}\right) e^{-\frac{x^2}{2}}.$$

By the expansion of $P_1(x)$ in a Fourier series we easily evaluate

$$d_n(t) = \int_{-\infty}^{\infty} e^{itx}\, d D_n(x)$$

and find

(10) $$d_n(t) \equiv -\frac{it}{\pi\sqrt{n}} \sum_{k=-\infty}^{\infty}{}' \frac{1}{ik} e^{-\frac{1}{2}(t+\pi\sqrt{n}k)^2},$$

the summation being performed for every integer $k \neq 0$.

Now let us consider the general case: $F(x) < (A;\, t_0)$. In analogy with (10) let us put

(11) $$d_n(t) \equiv -\frac{it}{t_0\sigma\sqrt{n}} \sum_{k=-\infty}^{\infty}{}' \frac{e^{-it_0 x_0 n k}}{ik} e^{-\frac{1}{2}(t+t_0\sigma\sqrt{n}k)^2},$$

x_0 being the least non-negative discontinuity point of $F(x)$. $D_n(x)$ being defined by

$$d_n(t) = \int_{-\infty}^{\infty} e^{itx}\, dD_n(x), \tag{12}$$

we find that

$$D_n(x) \equiv \frac{2\pi}{t_0 \sigma} \frac{1}{\sqrt{2\pi n}} P_1\left(\frac{(x - \xi_n)\, t_0\, \sigma \sqrt{n}}{2\pi}\right) e^{-\frac{x^2}{2}}. \tag{13}$$

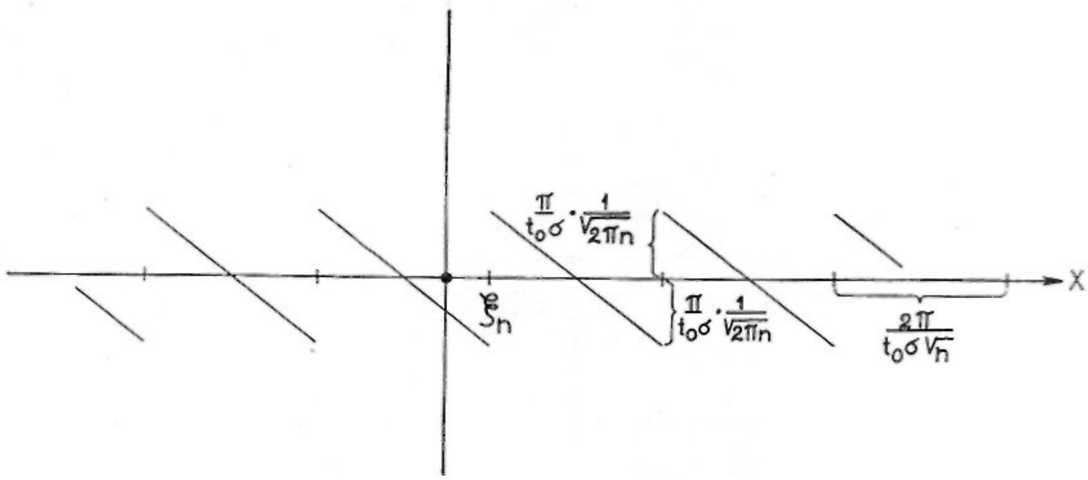

Fig. 3.

Let us put

$$\psi_n(x) \equiv P_1\left(\frac{(x - \xi_n)\, t_0\, \sigma \sqrt{n}}{2\pi}\right). \tag{14}$$

Thus $\psi_n(x)$ is a discontinuous periodic function with the period $\frac{2\pi}{t_0 \sigma \sqrt{n}}$ and the saltus 1. The translation ξ_n is determined as the least non-negative discontinuity point. It is found that

$$\xi_n = \frac{2\pi}{t_0 \sigma \sqrt{n}} \left\{ \frac{n x_0 t_0}{2\pi} - \left[\frac{n x_0 t_0}{2\pi}\right] \right\}. \tag{15}$$

The function $D_n(x) \cdot e^{\frac{x^2}{2}}$ is represented in fig. 3.

We can now state the following

Theorem 1. *Let $x_1, x_2, \ldots, x_n$ be a sequence of independent random variables with the same distribution function $F(x)$, the mean value zero, the dispersion $\sigma \neq 0$, the third moment α_3 and* *the finite absolute third moment* β_3. *Further let* $F(x) < (A;\ t_0)$. *Then*

$$(16)\quad F_n(x) = \Phi(x) + \frac{\alpha_3}{6\sigma^3\sqrt{2\pi n}}(1-x^2)\,e^{-\frac{x^2}{2}} + \\ + \frac{2\pi}{t_0\,\sigma\sqrt{2\pi n}}\,\psi_n(x)\,e^{-\frac{x^2}{2}} + o\left(\frac{1}{\sqrt{n}}\right)$$

as $n \to \infty$, $\psi_n(x)$ *being defined by* (14).

Remark. We observe that $P_1(x)$ is the same function that occurs in the EULER summation formula. Further we may notice that $D_n(x)$, apart from n, only depends on the two parameters σ and t_0, and on x_0 regarding the determination of an initial position.

In order to prove theorem 1 we apply the following

Theorem 2. *Let* A, T *and* ε *be arbitrary positive constants, let the distribution function* $V(x)$ *be purely discontinuous and* $G(x)$ *be a function of bounded variation* $(-\infty, \infty)$, $v(t)$ *and* $g(t)$ *the corresponding* FOURIER-STIELTJES *transforms, so that*

1. $G(-\infty) = 0$, $G(+\infty) = 1$.
2. *if* $G(x)$ *is discontinuous for* $x = x_\nu$, $(x_\nu < x_{\nu+1},\ \nu = 0, \pm 1, \pm 2, \ldots)$, *there exists a constant* $L > 0$ *such that* $Min\ (x_{\nu+1} - x_\nu) \geq L$,
3. $|G'(x)| \leq A < \infty$ *everywhere except when* $x = x_\nu$,
4. $V(x)$ *may only be discontinuous for* $x = x_\nu$,

$$(17)\qquad 5.\quad \int_{-T}^{T} \left|\frac{v(t) - g(t)}{t}\right| dt = \varepsilon.$$

Then to every number $k > 1$ *there correspond two finite positive numbers* $c(k)$ *and* $d(k)$, *only depending on* k, *such that*

$$(18)\qquad |V(x) - G(x)| \leq k \cdot \frac{\varepsilon}{2\pi} + c(k) \cdot \frac{A}{T},$$

provided that $T \cdot L \geq d(k)$.

Theorem 2 is analogous to a theorem that I have given before[6] and may be proved in the same manner. We apply theorem 2 to the proof of theorem 1, putting

[6] Cf. loc. cit.[2], theorem 1.

$$V(x) = F_n(x), \; G(x) = \Phi(x) + \frac{\alpha_3}{6\sigma^3\sqrt{2\pi n}}(1-x^2)e^{-\frac{x^2}{2}} + \\ + \frac{2\pi}{t_0\sigma\sqrt{2\pi n}}\psi_n(x)e^{-\frac{x^2}{2}},$$

$$v(t) = f_n(t), \; g(t) = e^{-\frac{t^2}{2}} - \frac{i\,\alpha_3}{6\sigma^3\sqrt{n}}t^3 e^{-\frac{t^2}{2}} + d_n(t),$$

$$A = \frac{\text{const.}}{\sqrt{n}}, \quad L = \frac{2\pi}{t_0\sigma\sqrt{n}}, \quad T = \text{const.}\cdot n.$$

By means of some expansions of $f_n(t)$ given earlier[7] it is possible to show that $\varepsilon = o\left(\frac{1}{\sqrt{n}}\right)$ as $n \to \infty$, ε being defined by (17). Hence theorem 1 is a consequence of theorem 2.

3. In this section it is no longer necessary that $F(x) < (A)$, but we suppose that the other conditions of theorem 1 are valid. By means of (3) and (16) we can now study the asymptotic behaviour of $\sqrt{n}\,|F_n(x) - \Phi(x)|$ and obtain the upper bound of this expression as $n \to \infty$. For the sake of simplicity we suppose $\alpha_3 = 0$.

If $F(x) < (A;\, t_0)$, let

$$d = \frac{2\pi}{t_0}. \tag{19}$$

By section 1 the distance between any two discontinuity points of $F(x)$ is an entire multiple of d and there is no quantity larger than d with this property.

By the combination of (3) and (16) the following theorem is obtained:

Theorem 3. *If the third moment* $\alpha_3 = 0$, *then*

$$\lim_{n\to\infty} \operatorname*{Max}_{-\infty<x<\infty} \sqrt{n}\,|F_n(x) - \Phi(x)| = \tag{20}$$

$$= \begin{cases} 0 & \text{if } F(x) \not< (A) \\ \dfrac{\pi}{t_0\sigma\sqrt{2\pi}} = \dfrac{d}{2\sigma\sqrt{2\pi}} & \text{》 } F(x) < (A;\, t_0) \end{cases}$$

[7] Cf. loc. cit.[2], lemma 3 and 4.

d being defined by (19).

$F(x)$ being subject to some further conditions, we can state:

Theorem 4. *Let $F(x)$ be symmetric*[8] *and continuous for $x=0$, then*

$$\lim_{n\to\infty} \operatorname*{Max}_{-\infty<x<\infty} \sqrt{n}\,|F_n(x)-\Phi(x)| \le \frac{1}{\sqrt{2\pi}}. \tag{21}$$

There is equality if and only if $F(x)$ is a symmetric Bernoulli *distribution, having the jump $\frac{1}{2}$ for $x=a$ and $x=-a$, a being a positive constant.*

Proof of theorem 4. Here $\alpha_3=0$. By theorem 3 it is sufficient to treat the case $F(x)<(A;\,t_0)$. We thus have to find an upper bound of

$$\frac{d}{2\sigma\sqrt{2\pi}}. \tag{22}$$

We may suppose on grounds of homogeneity that $d=1$. Thus under the given conditions $F(x)$ has the saltus $a_\nu \ge 0$ for $x=\pm\left(\nu+\frac{1}{2}\right)$, $(\nu=0,1,2,\ldots)$. Hence

$$\sum_{\nu=0}^{\infty} a_\nu=\frac{1}{2};\quad \sigma^2=2\sum_{\nu=0}^{\infty}\left(\nu+\frac{1}{2}\right)^2 a_\nu. \tag{23}$$

It is easily seen that the least possible value of σ is equal to $\frac{1}{2}$, $F(x)$ then being the symmetric Bernoulli distribution function with the saltus $\frac{1}{2}$ for $x=\pm\frac{1}{2}$. Hence $\frac{d}{2\sigma\sqrt{2\pi}}\le\frac{1}{\sqrt{2\pi}}$ and the theorem is proved.

Remark. If $F(x)$ does not satisfy any of the conditions of theorem 4, symmetry or continuity for $x=0$, it is possible to show that (22) is no longer bounded.

In a coming work I shall publish the complete proofs of

[8] A distribution function $F(x)$ is symmetric if $F(-x)=1-F(x)$.

theorems 1 and 2 and further state some asymptotic expansions which can be obtained if $F(x) < (A; t_0)$ and absolute moments of order > 3 are finite. At the same time I shall communicate some theorems on the dependence of the restterm on n and also on x, with applications.

Tryckt den 1 april 1943.

Uppsala 1943. Almqvist & Wiksells Boktryckeri A.-B.