

ON THE AVERAGE NUMBER OF REAL ROOTS OF A RANDOM ALGEBRAIC EQUATION

M. KAC

1. Introduction. Consider the algebraic equation

$$(1) \quad X_0 + X_1x + X_2x^2 + \cdots + X_{n-1}x^{n-1} = 0,$$

where the X 's are independent random variables assuming real values only, and denote by $N_n = N(X_0, \dots, X_{n-1})$ the number of real roots of (1). We want to determine the mean value (mathematical expectation = m.e.) of N_n when all X 's have the same normal distribution with density

$$(2) \quad e^{-u^2/\pi^{1/2}}.$$

This problem was treated by Littlewood and Offord¹ who also considered the cases when the X 's are uniformly distributed in $(-1, 1)$ or assume only the values $+1$ and -1 with equal probabilities. Littlewood and Offord obtain in each case the estimate

$$\text{m.e. } \{N_n\} \leq 25(\lg n)^2 + 12 \lg n, \quad n \geq 2000.$$

In our case of normally distributed X 's we shall be able to prove the exact formula

$$(3) \quad \text{m.e. } \{N_n\} = \frac{4}{\pi} \int_0^1 \frac{[1 - n^2[x^2(1 - x^2)/(1 - x^{2n})]^2]^{1/2}}{1 - x^2} dx$$

and then obtain the asymptotic relation

$$(4) \quad \text{m.e. } \{N_n\} \sim (2/\pi) \lg n$$

and the estimate

$$(5) \quad \text{m.e. } \{N_n\} \leq (2/\pi) \lg n + 14/\pi, \quad n \geq 2.$$

In case the X 's are not normally distributed (but all have the same distribution with standard deviation 1) one can still prove (4). The necessary limiting processes can then be carried out by using the central limit theorem of the calculus of probability and, as one may expect, the computations will be quite lengthy. On the other hand they will contribute relatively little to the general picture and, what

Presented to the Society, September 10, 1942; received by the editors July 2, 1942.

¹ J. London Math. Soc. vol. 13 (1938) pp. 288–295. No proofs are given in this paper, and the present author was unable to find them anywhere in print.

is worse, they may darken it by technicalities. We shall therefore assume in what follows that the X 's are normally distributed with density given by formula (2).

2. A formula for the number of real roots. Let $f(x)$ be a continuous function in $(-\infty, \infty)$ having a continuous first derivative $f'(x)$ and only a finite number of turning points in each finite interval.

Let $\psi_\epsilon(x)$ be 1 if $-\epsilon < x < \epsilon$ and 0 otherwise. We then have the following lemma.

LEMMA 1. *If neither a nor b is a zero of $f(x)$, then for sufficiently small ϵ 's*

$$(1/2\epsilon) \int_a^b \psi_\epsilon(f(x)) |f'(x)| dx$$

*is equal to the number of zeros of $f(x)$ inside the interval (a, b) . (Multiple roots are counted only once.)*²

We first notice that the set E_ϵ of those x 's for which $\psi_\epsilon(f(x)) = 1$ is an open set and is therefore a sum of disjoint open intervals I_1, I_2, I_3, \dots .

We choose ϵ small enough so that (a) no turning point of $f(x)$ in (a, b) lies in the strip $-\epsilon < y < \epsilon$ unless it happens to be at the same time a zero of $f(x)$; (b) no I_i includes either a or b .

Let I_1, I_2, \dots, I_r , say, be totally contained in (a, b) . Then, since it is easily seen that

$$\int_{I_i} |f'(x)| dx = 2\epsilon, \quad 1 \leq i \leq r,$$

we have

$$(1/2\epsilon) \int_a^b \psi_\epsilon(f(x)) |f'(x)| dx = (1/2\epsilon) \sum_{i=1}^r \int_{I_i} |f'(x)| dx = r.$$

This proves Lemma 1.

REMARK 1. An easy extension of the above reasoning gives for sufficiently small ϵ 's

$$(1/2\epsilon) \int_a^b \psi_\epsilon(f(x)) |f'(x)| dx = \text{number of zeros of } f(x) \text{ inside } (a, b) + \eta,$$

where $\eta = 0, 1/2, 1$ according as none, one or both of the numbers a, b are zeros of $f(x)$.

REMARK 2. If $f(x)$ is a polynomial then for sufficiently small ϵ 's

² If $f(x) \equiv 0$ we consider $f(x)$ as not having any roots.

$$(1/2\epsilon) \int_{-\infty}^{+\infty} \psi_\epsilon(f(x)) |f'(x)| dx$$

is equal to the number of real roots of $f(x)$. This is a trivial consequence of Lemma 1. One might notice that the limits of integration are really finite since $\psi_\epsilon(f(x)) = 0$ for sufficiently large $|x|$.

REMARK 3. The choice of ϵ in Remark 2 obviously depends on the coefficients of the polynomial. However, we can eliminate ϵ by re-stating the result of Remark 2 in the form:

$$\lim_{\epsilon \rightarrow 0} (1/2\epsilon) \int_{-\infty}^{+\infty} \psi_\epsilon(f(x)) |f'(x)| dx = \text{number of real roots of } f(x).$$

If we put now

$$f(x) = X_0 + X_1 x + \cdots + X_{n-1} x^{n-1},$$

we get

$$(6) \quad N_n = \lim_{\epsilon \rightarrow 0} (1/2\epsilon) \int_{-\infty}^{+\infty} \psi_\epsilon(f(x)) |f'(x)| dx.$$

We shall need this lemma in what follows.

LEMMA 2. *If $f(x)$ is a polynomial of degree $n-1$, then for every $\epsilon > 0$*

$$(1/2\epsilon) \int_{-\infty}^{+\infty} \psi(f(x)) |f'(x)| dx \leq 3n - 5.$$

In this case the set E (see proof of Lemma 1) is a sum of at most $2n-3$ open intervals. Indeed, each I_i contains either a real root of $f(x)$ or a turning point and there are at most $n-1$ real roots and at most $n-2$ turning points.

The proof of Lemma 2 follows now from the obvious remark that

$$\int_{I_i} |f'(x)| dx \leq 2\epsilon(m_i + 1), \quad 1 \leq i \leq 2n-3,$$

where m_i is the number of turning points in I_i .

3. Interchange of two limiting processes. We now reduce the computation of m.e. $\{N_n\}$ to the computation of m.e. $\{\psi_\epsilon(f(x)) |f'(x)|\}$ by means of this lemma.

LEMMA 3.

$$\text{m.e. } \{N_n\} = \lim_{\epsilon \rightarrow 0} (1/2\epsilon) \int_{-\infty}^{+\infty} \text{m.e. } \{\psi_\epsilon(f(x)) |f'(x)|\} dx.$$

Random variables can be considered as measurable functions defined on a set Ω with a completely additive Lebesgue measure. (The measure of Ω is 1.) Mathematical expectation (m.e.) is nothing but a Lebesgue integral with respect to that measure. Both N_n and $\psi_\epsilon(f(x))|f'(x)|$ are then measurable functions on Ω and they can be represented symbolically as

$$N_n(\mu) \quad \text{and} \quad g_\epsilon(x, \mu),$$

x and ϵ being real parameters. We also have

$$\begin{aligned} \text{m.e. } \{N_n\} &= \int_{\Omega} N_n(\mu) d\mu, \\ \text{m.e. } \{g_\epsilon(x, \mu)\} &= \int_{\Omega} g_\epsilon(x, \mu) d\mu, \end{aligned}$$

$d\mu$ indicating that integration is being performed with respect to the Lebesgue measure in Ω . We first notice that

$$(7) \quad \text{m.e. } \left\{ \int_{-\infty}^{+\infty} g_\epsilon(x, \mu) dx \right\} = \int_{-\infty}^{+\infty} \text{m.e. } \{g_\epsilon(x, \mu)\} dx.$$

This follows from Fubini's theorem if one notices that the integral with respect to dx is really an integral between finite limits (depending, of course, on ϵ and μ). Furthermore, by Lemma 2 for every $\epsilon > 0$

$$(1/2\epsilon) \int_{-\infty}^{+\infty} g_\epsilon(x, \mu) dx < 3n - 5,$$

and hence, by a well known theorem from Lebesgue's theory,

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \int_{\Omega} \left[(1/2\epsilon) \int_{-\infty}^{+\infty} g_\epsilon(x, \mu) dx \right] d\mu \\ = \int_{\Omega} \left[\lim_{\epsilon \rightarrow 0} (1/2\epsilon) \int_{-\infty}^{+\infty} g_\epsilon(x, \mu) dx \right] d\mu. \end{aligned}$$

This when combined with (6) and (7) completes the proof of Lemma 3.

4. A formula for m.e. $\{N_n\}$. In order to compute m.e. $\{\psi_\epsilon(f(x))|f'(x)|\}$ we need the following lemma.

LEMMA 4. *If $\alpha_0, \alpha_1, \dots, \alpha_{n-1}, \beta_0, \beta_1, \dots, \beta_{n-1}$ are real numbers, $\sum \alpha_i^2 = \alpha$, $\sum \beta_i^2 = \beta$, $\sum \alpha_i \beta_i = \gamma$ and if $\Delta = \alpha\beta - \gamma^2 > 0$, then the density of the joint distribution of $\alpha_0 X_0 + \dots + \alpha_{n-1} X_{n-1}$ and $\beta_0 X_0 + \dots + \beta_{n-1} X_{n-1}$ is equal to*

$$(1/\pi\Delta^{1/2}) \exp \{ - (\beta u^2 - 2\gamma uv + \alpha v^2)/\Delta \}.$$

This fact is well known.³

It also follows from well known facts that

$$\begin{aligned} \text{m.e. } & \{ \psi_\epsilon(\alpha_0 X_0 + \dots + \alpha_{n-1} X_{n-1}) \mid \beta_0 X_0 + \dots + \beta_{n-1} X_{n-1} \} \\ & = (1/\pi\Delta^{1/2}) \int_{-\infty}^{+\infty} \psi_\epsilon(u) \mid v \mid \exp \{ - (\beta u^2 - 2\gamma uv + \alpha v^2)/\Delta \} du dv. \end{aligned}$$

Let

$$\int_{-\infty}^{+\infty} \mid v \mid \exp \{ - (\beta u^2 - 2\gamma uv + \alpha v^2)/\Delta \} dv = F(u),$$

then $F(u)$ is a continuous function of u and we get

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} (1/2\epsilon) \text{ m.e. } & \{ \psi_\epsilon(\sum \alpha_i X_i) \mid \sum \beta_i X_i \} \\ & = \lim_{\epsilon \rightarrow 0} (1/2\epsilon\pi\Delta^{1/2}) \int_{-\epsilon}^{\epsilon} F(u) du = F(0)/\pi\Delta^{1/2}. \end{aligned}$$

But

$$F(0) = \int_{-\infty}^{+\infty} \mid v \mid \exp \{ - \alpha v^2/\Delta \} dv = \Delta/\alpha$$

and finally

$$\lim_{\epsilon \rightarrow 0} (1/2\epsilon) \text{ m.e. } \{ \psi_\epsilon(\sum \alpha_i X_i) \mid \sum \beta_i X_i \} = \Delta^{1/2}/\alpha.$$

If we now put $\alpha_0 = 1$, $\alpha_1 = x$, \dots , $\alpha_{n-1} = x^{n-1}$, $\beta_0 = 0$, $\beta_1 = 1$, $\beta_2 = 2x$, \dots , $\beta_{n-1} = (n-1)x^{n-2}$, we obtain by an elementary computation

$$\Delta = \frac{x^{4n} - n^2 x^{2(n+1)} + 2(n^2 - 1)x^{2n} - n^2 x^{2(n-1)} + 1}{(x^2 - 1)^4}$$

It is easy to prove that $\Delta > 0$ for every real x and by combining the considerations of this section with Lemma 3 we obtain

$$\begin{aligned} \text{m.e. } & \{ N_n \} \\ (8) \quad & = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{(x^{4n} - n^2 x^{2(n+1)} + 2(n^2 - 1)x^{2n} - n^2 x^{2(n-1)} + 1)^{1/2}}{(x^2 - 1)^2(1 + x^2 + x^4 + \dots + x^{2n-2})} dx. \end{aligned}$$

³ See for instance S. Bernstein, *Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes*, Math. Ann. vol. 97 (1927) pp. 2-59, in particular, chap. 3 (pp. 43-59). This chapter contains limit theorems by means of which one can handle the case of not normally distributed X 's. Lemma 4 is but a simple consequence of the main results of that chapter.

5. Estimates and the asymptotic formula. By an elementary transformation we get

$$\text{m.e. } \{N_n\} = \frac{4}{\pi} \int_0^1 \frac{(1 - h_n^2(x))^{1/2}}{1 - x^2} dx,$$

where

$$h_n(x) = \frac{nx^{n-1}(1 - x^2)}{1 - x^{2n}}.$$

It should be mentioned that the value of the integrand at $x=1$ is $(n^2 - 1/12)^{1/2}$. We have

$$1 - x^{2n} = (1 - x)(1 + x + x^2 + \dots + x^{2n-1}) < 2n(1 - x)$$

and therefore

$$\begin{aligned} h_n(x) &> x^{n-1}(1 + x)/2 \\ 1 - h_n^2(x) &< [1 - x^{n-1}(1 + x)/2][1 + x^{n-1}(1 + x)/2] \\ &< 2 - x^{n-1}(1 + x). \end{aligned}$$

Using the mean value theorem we get

$$\begin{aligned} 2 - x^{n-1}(1 + x) &= (1 - x)[\theta^{n-1} + (n - 1)\theta^{n-2}(1 + \theta)] \\ &< (2n - 1)(1 - x), \quad x < \theta < 1. \end{aligned}$$

Finally,

$$(1 - h_n^2(x))^{1/2}/(1 - x^2) < (2n - 1)^{1/2}/(1 - x)^{1/2}, \quad 0 \leq x < 1.$$

On the other hand

$$(1 - h_n^2(x))^{1/2}/(1 - x^2) \leq 1/(1 - x^2), \quad 0 \leq x < 1,$$

and we can write

$$\begin{aligned} \int_0^1 \frac{(1 - h_n^2(x))^{1/2}}{1 - x^2} dx &< \int_0^{1-1/n} \frac{dx}{1 - x^2} + \int_{1-1/n}^1 \frac{(2n - 1)^{1/2}}{(1 - x)^{1/2}} dx \\ &= (1/2) \lg n + (1/2) \lg (2 - 1/n) + 2(2 - 1/n)^{1/2} < (1/2) \lg n + 3 \cdot 5. \end{aligned}$$

Finally,

$$(9) \quad \text{m.e. } \{N_n\} < (2/\pi) \lg n + 14/\pi, \quad n \geq 2.$$

In order to obtain a lower estimate let ϵ and δ be arbitrary positive numbers less than 1. We have

$$\int_0^1 \frac{(1 - h_n^2(x))^{1/2}}{1 - x^2} dx > \int_0^{1-n^{\delta-1}} \frac{(1 - h_n^2(x))^{1/2}}{1 - x^2} dx.$$

But for $0 \leq x \leq 1 - n^{\delta-1}$

$$h_n(x) < nx^{n-1} \leq n(1 - n^{\delta-1})^{n-1}$$

and this last expression can be made smaller than $\epsilon^{1/2}$ if n is sufficiently large. Hence, for sufficiently large n ,

$$\begin{aligned} \int_0^1 \frac{(1 - h_n^2(x))^{1/2}}{1 - x^2} dx &> \int_0^{1-n^{\delta-1}} \frac{(1 - \epsilon)^{1/2}}{1 - x^2} dx \\ &> \frac{(1 - \epsilon)^{1/2}(1 - \delta)}{2} \lg n. \end{aligned}$$

Using (9) and the fact that ϵ and δ can be made arbitrarily small we obtain the asymptotic formula

$$(10) \quad \text{m.e. } \{N_n\} \sim (2/\pi) \lg n.$$

Let us add that it follows from (9) that the probability that (1) has more than $\sigma \lg n$ real roots is less than

$$2/\pi\sigma + 14/\pi\sigma \lg n.$$

This result is not trivial only in case $\sigma > 2/\pi$.

6. Final remarks. It is quite clear that the average number of real roots of (1) falling in the interval (a, b) is given by formula (8) if one replaces $-\infty$ and $+\infty$ by a and b , respectively. The proof of this statement does not differ from the one given above and one must only notice that the probability that either a or b is a root of (1) is 0.

One can also see almost immediately that if (a, b) does not contain either 1 or -1 the average number of real roots of (1) falling within (a, b) is $O(1)$. This means, roughly speaking, that most of the real roots of most of the equations cluster around 1 and -1 . The problem of the exact determination of the average distribution of real roots of (1) on the real axis will, of course, depend on a more delicate treatment of the integral (8). It would also be interesting to know the higher moments of N_n . The present line of attack would lead to very complicated integrals, but it may be hoped that some other approach will furnish more information about the distribution of the number of real roots of equation (1).