REAL ZEROS OF A RANDOM POLYNOMIAL
WITH LEGENDRE ELEMENTS

K. FARAHMAND
University of Ulster, Department of Mathematics
Jordanstown, Co. Antrim BT37 0QB, United Kingdom

(Received May, 1996; Revised December, 1996)

Let \( T_0^*(x), T_1^*(x), \ldots, T_n^*(x) \) be a sequence of normalized Legendre polynomials orthogonal with respect to the interval \((-1, 1)\). The asymptotic estimate of the expected number of real zeros of the random polynomial \( g_0 T_0^*(x) + g_1 T_1^*(x) + \ldots + g_n T_n^*(x) \) where \( g_j, j = 0, 1, \ldots, n \) are independent identically and normally distributed random variables with mean zero and variance one is known. The present paper considers the case when the means and variances of the coefficients are not all necessarily equal. It is shown that in general this expected number of real zeros is only dependent on variances and is independent of the means.

**Key words:** Orthogonal Polynomial, Weight Functions, Kac-Rice Formula, Probability Space, Jenser’s Theorem, Darboux-Christoffel Formula, Recurrence Formula.

**AMS subject classifications:** 60H99, 42BXX.

1. Introduction

Let \( \{g_j(\omega)\}_{j=0}^n \) be a sequence of independent normally distributed random variables defined on a probability space \((\Omega, \mathcal{A}, P_r)\). Let \( N_n(a, b) \) be the number of real zeros of \( P_n(x) \) in the interval \((a, b)\) where

\[
P_n(x) = P_n(x, \omega) = \sum_{j=0}^{n} g_j(\omega) T_j^*(x)
\]

in which

\[
T_j^*(x) = (\sqrt{j + 1/2}) T_j(x),
\]

and \( T_j(x) \) is a Legendre polynomial and therefore \( T_j^*(x) \) is a normalized Legendre polynomial orthogonal with respect to the weight function unity. For the case of identical normal standard distributed coefficients Das [3] shows that \( EN_n(-1, 1) \sim n/\sqrt{3} \) when \( n \) is sufficiently large. Recently in an interesting paper using a delicate method, Wilkins [12] shows that \( EN_n(-1, 1) = n/\sqrt{3} + o(n^\delta) \) for any positive \( \delta \).

Here we consider the case when the means and variances of the coefficients of
(1.1) are not all equal. We show that $EN_n(-1,1)$ is independent of $E(g_j)$ but dependent on $\text{var}(g_j)$. This is in complete contrast to the result obtained by Farahmand [5] for the random algebraic polynomial $\sum_{j=0}^{n} g_j(\omega)x^2$, who showed that $EN_n(-\infty,\infty)$ is independent of variances and dependent on the means. Nevertheless, as far as $EN_n$ is concerned there are similarities between polynomials of type (1.1) and random trigonometric polynomials $\sum_{j=0}^{n} g_j(\omega)\cos jx$, see for example [4], in that both have $O(n)$ number of real zeros and both are effected by the variance of the coefficients and not by their means. A survey of the earlier works together with comprehensive references on the subject can be found in Bharucha-Reid and Sambandham [1].

We prove the following theorems:

**Theorem 1:** If the coefficients of $\{g_j\}_{j=0}^{n}$ of $P_n(x)$ have means and variances $\mu_1$ and $\sigma_1^2 > 0$ for $0 \leq j < n'$ and $\mu_2$ and $\sigma_2^2 > 0$ for $n' < j \leq n$, respectively, then for all sufficiently large $n$, the mathematical expectation of the number of real zeros of $P_n(x)$ satisfies

$$EN_n(-1,1) \sim \frac{\left(\frac{n'+1}{2}\right)^{3}(2n'+3)^{1/2}(\frac{n'+1}{2})^{1/2} + \frac{n^2\sigma_2^2}{n'^3\sigma_1^2}}{\sqrt{3}\left(\frac{n'+1}{2}\right)^{3}(2n'+3)^{1/2}(\frac{n'+1}{2})^{1/2} + \frac{n\sigma_2^2}{n'^3\sigma_1^2}}^{1/2}.$$

Although Theorem 1 is of interest in that it allows any integer values for $0 \leq n' \leq n$, it is of interest to study $EN_n(-1,1)$ for the special cases whether or not $n'/n$ tends to zero. The following theorems, which are in fact the corollaries of Theorem 1, give these results.

**Theorem 2:** If the coefficients $g_j$'s are distributed according to the assumptions of Theorem 1, then for $n' \to \infty$ as $n \to \infty$,

$$EN_n(-1,1) \sim \frac{n'^3\sigma_1^2 + (n^3 - n'^3)\sigma_2^2}{\sqrt{3}(n'^3\sigma_1^2 + (n^3 - n'^3)\sigma_2^2)}^{1/2}.$$

**Theorem 3:** Under the assumptions of Theorem 1, for $n'/n \to 0$ as $n \to \infty$,

$$EN_n(-1,1) \sim \frac{n}{\sqrt{3}}.$$

The result in Theorem 3 does, indeed, correspond to that of Das [3], and therefore it shows that Das' result remains valid for polynomials with non-identical distributed coefficients. These non-identical cases, of interest in their own right, are important as they lead to the expected number of crossings of two polynomials with different degrees. Let

$$F_n(x) = \sum_{j=0}^{n} f_j(\omega)T_j^*(x) \quad \text{and} \quad Q_n(x) = \sum_{j=0}^{n'} q_j(\omega)T_j^*(x).$$

Then, for $n > n'$, the expected number of real zeros of polynomial

$$F_n(x) - Q_n(x) = \sum_{j=0}^{n'} \{f_j(\omega) - q_j(\omega)\}T_j^*(x) + \sum_{j=n'}^{n} f_j(\omega)T_j^*(x),$$

can serve as the expected number of crossings of $F_n(x)$ by $Q_n(x)$. On the other hand, $F_n(x) - Q_n(x)$ can be represented as $\sum_{j=0}^{n} g_j(\omega)T_j^*(x)$ where $g_j(\omega) = f_j(\omega) - q_j(\omega)$ for $0 < j \leq n'$ and $g_j(\omega) = q_j(\omega)$ for $n' < j \leq n$, which is, indeed, in the form of $F_n(x)$ studied in Theorem 1-3.
2. A Formula for the Expected Number of Crossings

Let \( \Phi(t) \) and \( \phi(t) \) be the distribution and density functions of a normal standard random variable, respectively. Then from Cramér and Leadbetter [2, page 285] we have,

\[
EN_n(a, b) = \int_a^b \frac{B(1 - \theta^2)^{1/2}}{A} \phi \left( \frac{x}{A} \right) \left[ 2\phi(\eta) + \eta \{2\Phi(\eta) - 1\} \right] dx, \tag{2.1}
\]

where

\[
A^2 = \text{var}\{ P_n(x) \}, \quad B^2 = \text{var}\{ P'_n(x) \},
\]

\[
\theta = \frac{\text{cov}\{ P_n(x), P'_n(x) \}}{AB}, \quad \lambda_1 = E\{ P_n(x) \},
\]

\[
\lambda_2 = E\{ P'_n(x) \} \quad \text{and} \quad \eta = \frac{\lambda_2 - B\theta\lambda_1 / A}{B(1 - \theta^2)^{1/2}}.
\]

Let \( C = \text{cov}\{ P_n(x), P'_n(x) \} \), \( \Delta^2 = A^2B^2 - C^2 \) and \( \text{erf}(x) = \int_0^x \exp(-t^2) dt \), then from (2.1) we can write the extension of a formula for the expected number of real zeros obtained by Kac [7] and Rice [8] as

\[
EN_n(a, b) = I_1(a, b) + I_2(a, b), \tag{2.2}
\]

where

\[
I_1(a, b) = \int_a^b \frac{\Delta}{\pi A^2} \exp \left( - \frac{A^2\lambda_2^2 - 2C\lambda_1\lambda_2 + B^2\lambda_1^2}{2\Delta^2} \right) dx \tag{2.3}
\]

and

\[
I_2(a, b) = \int_a^b \frac{\sqrt{2} | A^2\lambda_2 - C\lambda_1 |}{\pi A^3} \exp \left( - \frac{\lambda_1^2}{2A^2} \right) \text{erf} \left( \frac{A^2\lambda_2 - C\lambda_1}{\sqrt{2}A\Delta} \right) dx. \tag{2.4}
\]

Now let \( T_k^{(i)}(x) \), \( k = 0, 1, \ldots, n; i = 1, 2, 3 \) represent the \( i \)th derivative of \( T_k(x) \) with respect to \( x \) and set

\[
R_k^{(i)}(x) = T_k^{(i)}(x)T_{k+1}^{(i)}(x) - T_k^{(i)}(x)T_{k+1}^{(i)}(x), \quad i = 0, 1, 2, 3; j = 0, 1.
\]

Then from the Darboux-Christoffel formula [6, page 1024] putting \( d_k = (k + 1)(2k + 3)^{1/2} / 2(2k + 1)^{1/2} \) we can write

\[
A^2 = \sigma_1^2 \sum_{j=0}^{n'} T_j^*(x) + \sigma_2^2 \sum_{j=n'+1}^{n} T_j^*(x) = \sigma_1^2 d_{n'} R_{n'}^{10}(x) + \sigma_2^2 \{ d_n R_{n}^{10}(x) - d_{n'} R_{n'}^{10}(x) \}. \tag{2.5}
\]

Similarly we obtain

\[
B^2 = \sigma_1^2 \left\{ \frac{d_{n'} R_{n}^{30}(x) + d_{n'} R_{n}^{21}(x)}{6} \right\} + \sigma_2^2 \left\{ \frac{d_n R_{n}^{30}(x) + d_n R_{n}^{21}(x)}{2} - \frac{d_{n'} R_{n'}^{30}(x) - d_{n'} R_{n'}^{21}(x)}{6} \right\}. \tag{2.6}
\]
and
\[ C = \frac{\sigma^2_d}{2} d' n R^{20}_n(x) + \frac{\sigma^2_d}{2} \left( d_n R^{20}_n(x) - d_n R^{20}_n(x) \right). \] (2.7)

At this stage we abandoned the calculations of \( \lambda_1 \) and \( \lambda_2 \) as only an upper limit for these are required. We proceed to evaluate these upper limits as well as the dominant terms for \( A^2, B^2 \) and \( C \) in the following section.

3. Approximations

In order to find the asymptotic value of \( EN_n(-1,1) \) we let \( \epsilon = n^{-1/4} \) and we first evaluate \( EN_n(-1 + \epsilon, 1 - \epsilon) \). To this end, we use the Kac-Rice formula (2.2)-(2.4), and therefore, we need only to estimate the terms involved in (2.5)-(2.7) outside the \( \epsilon \)-neighborhoods of \( -1 \) and \( 1 \). For the intervals \( (-1, -1 + \epsilon) \) and \( (1, 1 - \epsilon) \) we need to modify Dunnage’s [4] approach which is based on Jensen’s theorem [9, page 332] or [11, page 125].

For Legendre polynomials we have the following recurrence formula, see for example [6, page 1026],

\[ T_{k+1}(x) = \frac{k+1}{1-x^2} \{ T_k(x) - x T_{k+1}(x) \}. \] (3.1)

Rewriting (3.1) for \( k \) and using the following recurrence formula also valid for any Legendre polynomial

\[ kT_{k-1}(x) = (2k+1)x T_k(x) - T_{k+1}(x). \]

we obtain

\[ T'_k(x) = \frac{k+1}{1-x^2} \{ x T_k(x) - T_{k+1}(x) \}. \] (3.2)

Therefore from (3.1) and (3.2) we can write

\[ R_{k}^{10}(x) = T_{k+1}(x) T_k(x) - T_{k+1}(x) T'_k(x) \]

\[ = \frac{k+1}{1-x^2} \left\{ T_{k+1}^2(x) + T_{k}^2(x) - 2x T_k(x) T_{k+1}(x) \right\}. \] (3.3)

Now in order to estimate \( R_{n}^{10}(x) \) and \( R_{k}^{10}(x) \) in (3.3) we assume \( x \in (-1 + \epsilon, 1 - \epsilon) \). Then from [10, page 195, Theorem 8.21.5] (see also [12]), letting \( x = \cos \gamma \) we have

\[ T_k(\cos \gamma) = \sqrt{\frac{2}{k \sin \gamma}} \sum_{\nu=0}^{p-1} \frac{k!(\nu + 1/2)^{2}}{\pi \nu!(2 \sin \gamma)^{2}} \frac{1}{T(k + \nu + 3/2)} \]

\[ + O(k \sin \gamma)^{-1/2}. \]

For our purpose of estimating \( R_{k}^{10}(x) \) we need only to put \( p = 1 \) which gives

\[ T_k(\cos \gamma) = \sqrt{\frac{2}{k \pi (1 - x^2)^{1/2}}} \cos \{(k + 1/2) - \pi/4\} + O(k \sin \gamma)^{-3/2}. \]

Therefore the following term of \( R_{k}^{10}(x) \) which appeared in (3.3) can be written as

\[ T_{k}^2(x) + T_{k+1}^2(x) - 2x T_k(x) T_{k+1}(x) \]
\[ \frac{2}{k \pi \sqrt{1 - x^2}} \left( \cos^2 \{(k + 1/2)\gamma - \pi/4\} + \cos^2 \{(k + 3/2)\gamma - \pi/4\} \right) - 2 \cos \gamma \cos \{(k + 1/2)\gamma - \pi/4\} \cos \{(k + 3/2)\gamma - \pi/4\} + O\left(\frac{1}{k^2(1 - x^2)}\right) \]

\[ = \frac{2\sqrt{1 - x^2}}{k \pi} + O\left(\frac{1}{k^2(1 - x^2)}\right). \tag{3.4} \]

Hence from (2.5), (3.3) and (3.4) letting

\[ c_k = \frac{2(k + 1)d_k}{k} = \frac{(k + 1)^2(2k + 3)^{1/2}}{k(2k + 1)^{1/2}}, \]

we get

\[ A^2 = \frac{1}{\pi(1 - x^2)^{1/2}} \left\{ c_n' (\sigma_1^2 - \sigma_2^2) + \sigma_2^2 c_n \right\} + O\left(1 - x^2\right)^{-2} \left(\frac{1}{n^2} + \frac{1}{n} \right). \tag{3.5} \]

It is a well known property of any Legendre polynomial that it satisfies the following differential equation

\[ T''_k(x) = \frac{2x T'_k(x) - k(k + 1)T_k(x)}{1 - x^2}. \]

This formula and its equivalent written for \( k + 1 \) gives

\[ R_{k}^{21}(x) = -\frac{(k + 1)\{k R_{k}^{01}(x) + 2T_{k+1}(x)T'_k(x)\}}{1 - x^2} \tag{3.6} \]

and

\[ R_{k}^{20}(x) = \frac{2x R_{k}^{10}(x) - 2(k + 1)T_k(x)T_{k+1}(x)}{1 - x^2}. \tag{3.7} \]

Now since from the first theorem of Stielzer [10, page 197] \( T_k(x) = O\{k^{-1/2}(1 - x^2)^{-1/4}\} \), and therefore by (3.1), \( T'_k(x) = O\{k^{1/2}(1 - x^2)^{-5/4}\} \) from (3.6) and (3.7) we obtain

\[ R_{k}^{21}(x) = \frac{k(k + 1)R_{k}^{10}(x)}{1 - x^2} + O\left(\frac{k}{(1 - x^2)^{5/2}}\right) \tag{3.8} \]

and

\[ R_{k}^{20}(x) = \frac{2x R_{k}^{10}(x)}{1 - x^2} + O\left(\frac{1}{(1 - x^2)^{3/2}}\right). \tag{3.9} \]

Similarly differentiating (3.7) and using (3.6) yields

\[ R_{k}^{30}(x) = \frac{8x^2/(1 - x^2) - n(n + 1)\}R_{k}^{10}(x)}{1 - x^2} + O\left(\frac{1}{(1 - x^2)^{5/2}}\right). \tag{3.10} \]

Now (3.8)-(3.10) are sufficient to evaluate (2.6) and (2.7) as

\[ B^2 = \frac{(\sigma_1^2 - \sigma_2^2)h_n' + \sigma_2^2 h_n}{3\pi(1 - x^2)^{3/2}} + O\left(\frac{n^2}{(1 - x^2)^{5/2}}\right) \tag{3.11} \]

and

\[ C = O\left(\frac{n}{(1 - x^2)^{3/2}}\right) \tag{3.12} \]

where

\[ h_k = 2(k + 1)^2 d_k = \frac{(k + 1)^3(2k + 3)^{1/2}}{(2k + 1)^{1/2}}. \]
For $\lambda_1$ and $\lambda_2$, as we will see, it would be sufficient to obtain only their order. To this end, let $m = \lfloor n^{2/3} \rfloor$ be the integer part of $n^{2/3}$, since $|T_n(x)| < 4n^{-1/2}(1 - x^2)^{-1/4}$ and

$$\sum_{j=0}^{k} (2j + 1)T_j(x) = (k + 1)\frac{T_k(x) - T_{k+1}(x)}{1 - x}, \quad (3.13)$$

we can write

$$\sum_{j=0}^{n} T_j^*(x) = O\left\{ \sum_{j=0}^{m} (j + 1/2)T_j(x) + \sum_{j=m+1}^{n} \frac{j + 1}{(m + 1/2)^{1/2}}T_j(x) \right\}$$

$$= O\left\{ (m + 1)\frac{|T_m(x) - T_{m+1}(x)|}{1 - x} + (n + 1)\frac{|T_n(x) - T_{n+1}(x)|}{\sqrt{m(1 - x)}} \right\}$$

$$= O\left\{ \frac{n^{2/3}|T_{n+1}(x)|}{1 - x} \right\} = O\left\{ \frac{n^{1/6}}{(1 - x)(1 - x^2)^{1/2}} \right\}. \quad (3.14)$$

Using (3.13), similarly we can obtain

$$\sum_{j=0}^{n} T_j^*(x) = O\left\{ \sum_{j=0}^{n} \frac{(j + 1/2)^{1/2}(j + 1)}{(2j + 1)(1 - x^2)}T_j(x) \right\}$$

$$= O\left\{ n^{1/2}\frac{n^{3/2}|T_n(x)|}{(1 - x)(1 - x^2)^{5/4}} \right\} = O\left\{ \frac{n}{(1 - x)(1 - x^2)^{5/4}} \right\}. \quad (3.15)$$

### 4. Proofs of Theorems

We first use the Kac-Rice formula (2.4) for $EN_n(-1 + \epsilon, 1 - \epsilon)$. We will see that this interval yields the main contribution to the expected number of real zeros. The expected number of real zeros outside this interval, which, it so happens, are negligible, is estimated by using an application of Jensen’s theorem.

From (2.3), (3.5), (3.11)-(3.15) we can easily write,

$$I_1(-1 + \epsilon, 1 - \epsilon) \sim \frac{\{(\sigma_1^2 - \sigma_2^2)h_{n'} + \sigma_2^2 h_n\}^{1/2}}{\pi \sqrt{3}\{(\sigma_1^2 - \sigma_2^2)c_{n'} + \sigma_2^2 c_n\}^{1/2}} \int_{-1 + \epsilon}^{1 - \epsilon} \frac{dx}{(1 - x^2)^{1/2}}$$

$$\sim \frac{\{(\sigma_1^2 - \sigma_2^2)h_{n'} + \sigma_2^2 h_n\}^{1/2}}{\sqrt{3}\{(\sigma_1^2 - \sigma_2^2)c_{n'} + \sigma_2^2 c_n\}^{1/2}}. \quad (4.1)$$

Also from (2.4), (3.5), (3.11)-(3.15) we can easily show
\[ I_2(-1 + \epsilon, 1 - \epsilon) = O(n^{1/2}). \]  

(4.2)

Now we show that the expected number of real zeros outside the interval \((-1 + \epsilon, 1 - \epsilon)\) is small compared to (4.1). To this end, let \(N(r)\) denote the number of zeros of \(P_n(x)\) in the circle \(|x - 1| \leq r\). Since an upper bound for \(N(\epsilon)\) could serve as an upper bound for the number of zeros in the interval \((1 - \epsilon, 1)\), we confine ourselves to \(N(\epsilon)\). The interval \((-1, -1 + \epsilon)\) can be treated exactly the same way to give the same result. It follows from Jensen’s theorem [11, page 332] or [9, page 125] that

\[ N(\epsilon) \leq (2\pi \log 2)^{-1} \int_0^{2\pi} \log \left( \frac{P_n(1 + \epsilon e^{ix}, \omega)}{P_n(1, \omega)} \right) dx. \]  

(4.3)

Now we use the identity

\[ T_n(z) = \frac{1}{\pi} \int_0^\pi \{z + i(1 - z^2)^{1/2}\cos \theta\}^n d\theta \]

for all sufficiently large \(n\), to obtain

\[ T_n(1 + \epsilon e^{ix}) < (1 + 3\epsilon)^n < \exp(3n\epsilon). \]

Therefore, since by Schwarts inequality

\[ \sum_{j=0}^{n} (j+1/2)^{1/2} \leq \left( n+1 \right) \sum_{j=0}^{n} (j+1/2)^{1/2} < n^{3/2}, \]

for sufficiently large \(n\) we can write

\[ P_n(1 + \epsilon e^{ix}) < n^{3/2} \exp(3n\epsilon) \max_{0 \leq j \leq n} |g_j|. \]

(4.4)

Now since \(g_j, j = 0, 1, 2, \ldots, n\) has a normal distribution,

\[ \Pr(\max_{0 \leq j \leq n} |g_j| > n) < n' \left( \frac{2}{\pi \sigma_1^2} \right)^{1/2} \int_0^{\infty} \exp \left\{ - \frac{(1 - \mu_1)^2}{2\sigma_1^2} \right\} dt \]

\[ + (n - n') \left( \frac{2}{\pi \sigma_2^2} \right)^{1/2} \int_0^{\infty} \exp \left\{ - \frac{(t - \mu_2)^2}{2\sigma_2^2} \right\} dt \]

\[ < \frac{2\sqrt{n' \sigma_1}}{\sqrt{\pi(n - \mu_1)}} \exp \left\{ - \frac{(n - \mu_1)^2}{2\sigma_1^2} \right\} + \frac{2(n-n') \sqrt{2\sigma_2}}{\sqrt{\pi(n - \mu_2)}} \exp \left\{ - \frac{(n - \mu_2)^2}{2\sigma_2^2} \right\} \]

\[ < 4 \exp \left\{ - \frac{(n - \mu)^2}{2\sigma^2} \right\}, \]  

(4.5)

where \(\mu = \max\{\mu_1, \mu_2\}\) and \(\sigma^2 = \min\{\sigma_1^2, \sigma_2^2\}\). Hence from (4.4) and (4.5) outside sample functions in an \(\omega\)-set of measure not exceeding \(4 \exp\{ - (n - \mu)^2/2\sigma^2 \}\),

\[ P_n(1 + \epsilon e^{ix}) < n^{5/2} \exp(3n\epsilon). \]  

(4.6)

Moreover, since \(T_j(1) = 1, j = 0, 1, 2, \ldots, n\), it follows that the distribution of \(P_n(1)\) is normal with mean

\[ m = \mu_1 \sum_{j=0}^{n'} (j+1/2)^{1/2} + \mu_2 \sum_{j=n'+1}^{n} (j+1/2)^{1/2} \]

and variance
\[ s^2 = (\sigma_1^2 - \sigma_2^2) \frac{n'(n' + 2)}{2} + \sigma_2^2 \frac{n(n + 2)}{2}. \]

Therefore,
\[
Pr\{ | P_n(1) | < 1 \} \leq \frac{1}{s\sqrt{2\pi}} \int_{-1}^{1} \exp \left\{ - \frac{(t - m)^2}{2s^2} \right\} dt < \frac{2}{\sqrt{\pi s^2}}.
\]

(4.7)

However from (4.4), (4.6), (4.7) and except for sample functions in an \( \omega \)-set of measure not exceeding \( \sqrt{2/\pi s^2 + 4\exp\{-(n-\mu)^2/2\sigma^2\} < (4/n)\max\{\sigma_1^2, \sigma_2^2\}}, \)
\[
N(\epsilon) \leq \frac{(5/2)\log n + 3n\epsilon}{\log 2},
\]
which gives \( O(n\epsilon + \log n) \) as the upper bound for \( EN(\epsilon) \). This upper limit is smaller than the error term involved in (4.1) and (4.2). Therefore the proof of Theorem 1 follows.

References