THE EXPECTED NUMBER OF ZEROS OF A STATIONARY GAUSSIAN PROCESS[1]

By N. Donald Ylvisaker

University of Washington

1. Introduction. It is assumed throughout that $\{X(t),\ t \varepsilon [0, 1]\}$ is a real separable stationary Gaussian process with mean function zero, covariance function ρ, $\rho(0) = 1$, and having continuous sample paths. It follows that there is a normalized (spectral) distribution function F such that $F(u) = 1 - F(-u)$ at points of continuity u and for which

$$EX(s)X(t) = \rho(s - t) = \int_{-\infty}^{\infty} \cos (s - t)\lambda \, dF(\lambda), \qquad s, t \,\varepsilon\, [0, 1].$$

We will use freely the equivalence of the two conditions: ρ'' exists continuous at the origin and $\int_{-\infty}^{\infty} \lambda^2 \, dF(\lambda) < \infty$; further, that these conditions imply $X(\cdot)$ is absolutely continuous a.s. (cf [3] p. 536).

Let N be the number of zeros of $X(\cdot)$. It is shown here that

(a) $X(\cdot)$ has a.s. no tangential zeros,

(b) $EN = (1/\pi)(-\rho''(0))^{\frac{1}{2}}$ if $\rho''(0)$ exists,
$= +\infty$ if not.

The formula for EN goes back to Rice [8] who obtains it under the assumption that F has finitely many points of increase. Ivanov [5] proves that EN is given above when $\rho^{(iv)}(0)$ exists (equivalently $\int_{-\infty}^{\infty} \lambda^4 \, dF(\lambda) < \infty$), while Bulinskaya [2] shows that (a) and (b) hold provided $X(\cdot)$ has a continuous derivative (for which the best sufficient condition known is $\int_{-\infty}^{\infty} \lambda^2 \log (1 + |\lambda|)^{1+\delta} \, dF(\lambda) < \infty$, $\delta > 0$). The result given here however is not unexpected (see for example [4], p. 273).

Most previous work in this area has followed Kac who in [6] devised a procedure for counting zeros; inasmuch as this procedure is designed to account for tangencies when there are none, we adopt a different, and perhaps simpler, approach to the counting. A suitable modification of what we do will produce the expected number of a's, $a \neq 0$, for the same processes and indications are that it can also be extended to handle zeros of nonstationary Gaussian processes with nonzero mean functions. Leadbetter and Cryer [7] have found the corresponding formula, applicable when the process has a continuous derivative, and they point out that this is the general situation for arbitrary barriers and Gaussian processes. For corresponding nonnormal problems it will be seen that the main calculations involve the bivariate distributions of the process with tangencies requiring special attention.

Received 28 October 1964.

[1] Research supported in part by the National Science Foundation under grants G-19645 and G-25211.

2. Zeros and expectations. If f is a continuous function on the unit interval, we say f has a crossing zero at $t_0 \varepsilon (0, 1)$ provided every neighborhood of t_0 contains points t_1 and t_2 with $f(t_1)f(t_2) < 0$; we say f has a tangential zero at $t_0 \varepsilon (0, 1)$ provided $f(t_0) = 0$ and there is a neighborhood of t_0 on which f has a constant sign. Let $C(T)$ be the number of crossing (tangential) zeros of f in $(0, 1)$ so that the number of zeros of f in $(0, 1)$ is $C + T$.

For the process $X(\cdot)$ we show $T = 0$ a.s., this is accomplished by referring tangencies to the continuity of the variables $\sup_{(a,b)} X(\cdot)$. If $\{t_j\}$ is a countable dense set in $(0, 1)$ and if $\{\epsilon_n\}$ is a sequence of positive numbers with limit 0, we note that

$$\{x(\cdot) \mid x(\cdot) \text{ has a tangential zero from below in } (0, 1)\} \subset \bigcup_{j,n} \{x(\cdot) \mid \sup_{(t_j-\epsilon_n, t_j+\epsilon_n)} x(\cdot) = 0\}.$$

LEMMA 1. $\sup_{(a,b)} X(\cdot)$ *has a continuous distribution for any interval* $(a, b) \subset (0, 1)$.

PROOF. If $\{T_n\}$ is an increasing sequence of finite sets in (a, b) with $\bigcup_n T_n$ dense in (a, b), $\max_{T_n} X(\cdot) \rightarrow_{\text{a.s.}} \sup_{(a,b)} X(\cdot)$. Now $\max_{T_n} X(\cdot)$ has a density of the form $\phi \cdot G_n$ where ϕ is the standard normal density and where $G_n(u)$ is a sum of terms of the form $P[X_1 \leqq u, \cdots, X_k \leqq u \mid X_0 = u]$. These conditional probabilities are in fact nondecreasing in u. To see this, suppose $(X_0, X_1, \cdots, X_k)$ is multivariate normal with $EX_i = 0, EX_i^2 = 1, i = 0, 1, \cdots, k$. For convenience, suppose $X_0, X_1, \cdots, X_r$ is a maximal linearly independent subset and that $X_i = \sum_{j=0}^{r} \theta_{ij}X_j, i = r + 1, \cdots, k$. If $EX_iX_j = \sigma_{ij}$ for $i, j = 0, 1, \cdots, r$, the conditional distribution of $(X_1, \cdots, X_r)$ given $X_0 = u$ is multivariate normal with mean $(\sigma_{01}u, \cdots, \sigma_{0r}u)$ and covariance matrix of (i, j)th entry $\sigma_{ij} - \sigma_{0i}\sigma_{0j}$. Now $P[X_1 \leqq u, \cdots, X_k \leqq u \mid X_0 = u]$ is an integral of the corresponding density over the set

$$[x_1 \leqq u, \cdots, x_r \leqq u, \textstyle\sum_{j=1}^{r} \theta_{ij}x_j \leqq u(1 - \theta_{i0}), i = r + 1, \cdots, k].$$

Centering the density by letting $y_i = x_i - \sigma_{0i}u$, it is an integral of a density independent of u over the set

$$[y_1 \leqq u(1 - \sigma_{01}), \cdots, y_r \leqq u(1 - \sigma_{0r}), \textstyle\sum_{j=1}^{r} \theta_{ij}y_j \leqq u(1 - \sum_{j=0}^{r} \theta_{ij}\sigma_{0j}), \quad i = r + 1, \cdots, k].$$

We see here that the coefficients of u are all nonnegative for, in particular, $\sum_{j=0}^{r} \theta_{ij}\sigma_{0j} = \sum_{j=0}^{r} \theta_{ij}EX_0X_j = EX_0X_i \leqq 1, i = r + 1, \cdots, k$. Thus the conditional probabilities in question are nondecreasing in u as is the function G_n. Now in order that the distribution function of $\sup_{(a,b)} X(\cdot)$ have a mass point at u_0 say, it is necessary that the sequence $\{G_n(u_0 + \epsilon)\}$ be unbounded, $\epsilon > 0$. However this cannot be so for

$$\left(\int_{u_0+\epsilon}^{\infty} \phi(u)\, du\right)G_n(u_0 + \epsilon) \leqq \int_{u_0+\epsilon}^{\infty} \phi(u)G_n(u)\, du \leqq 1.$$

Although it is not needed here, it does follow from the nature of the densities involved, that $\sup_{(a,b)} X(\cdot)$ is absolutely continuous with density $\phi \cdot G$, $G = \lim_n G_n$ and nondecreasing.

We now suppose f is a continuous function on [0, 1] for which $f(k2^{-n}) \neq 0$, $k = 0, 1, \cdots, 2^n$, $n = 1, 2, \cdots$. If $f(t_1)f(t_2) < 0$ for $t_1 < t_2$, then f has at least one crossing zero in (t_1, t_2). Consider the auxiliary variables

$$\begin{aligned} U_{nk} &= 1 \quad \text{if } f((k-1)2^{-n})f(k2^{-n}) < 0, \\ &= 0 \quad \text{otherwise}, \qquad k = 1, 2, \cdots, 2^n, \\ Z_n &= \textstyle\sum_{k=1}^{2^n} U_{nk}, \qquad n = 1, 2, \cdots. \end{aligned}$$

$\{Z_n\}$ is a nondecreasing sequence so let $Z = \lim_n Z_n$. As noted above $Z_n \leqq C$ and therefore $Z \leqq C$.

LEMMA 2. $Z = C$ *(both sides may be infinite).*

PROOF. If C is finite, the crossing zeros are separated and so are counted by some Z_n, $Z = C$. If moreover Z is finite, then the crossing intervals $((k-1)2^{-n}, k2^{-n})$ counted by $Z_n = Z$ can be separated for n sufficiently large and f must be of constant sign on the remaining noncrossing intervals. Letting $n \to \infty$ we find points $0 = t_0 < t_1 < \cdots < t_{Z+1} = 1$ such that f is of constant sign on $[t_i, t_{i+1}]$, $i = 0, 1, \cdots, Z$. Thus, C is finite.

For the process $X(\cdot)$ the sample functions are a.s. different from zero at all points of the form $k2^{-n}$, therefore $\{Z_n\}$ is a.s. nondecreasing with limit C. We find the expectations ([1], p. 43)

$$EU_{nk} = (1/\pi) \text{ arc cos } \rho(2^{-n}), \qquad EZ_n = (2^n/\pi) \text{ arc cos } \rho(2^{-n}).$$

LEMMA 3. $(2^n/\pi)$ arc cos $\rho(2^{-n})$ *has a finite limit if and only if* $\rho''(0)$ *exists, in this case* $(2^n/\pi)$ arc cos $\rho(2^{-n}) \to (1/\pi)(-\rho''(0))^{\frac{1}{2}}$.

PROOF. Suppose $(2^n/\pi)$ arc cos $\rho(2^{-n})$ has a finite limit. Since

$$\text{arc cos } \rho(2^{-n}) \geqq \{[1 - \rho(2^{-n})]/[1 + \rho(2^{-n})]\}^{\frac{1}{2}}$$

it follows that $[1 - \rho(2^{-n})]/[1 + \rho(2^{-n})] = O(2^{-2n})$ or $1 - \rho(2^{-n}) = O(2^{-2n})$. Consequently, each of the following is bounded in n:

(i) $2^{2n+1} \int_0^{2^{n/4}} (1 - \cos \lambda 2^{-n})\, dF(\lambda) + 2^{2n+1} \int_{2^{n/4}}^{\infty} (1 - \cos \lambda 2^{-n})\, dF(\lambda)$,

(ii) $2^{2n+1} \int_0^{2^{n/4}} (\lambda^2 2^{-(2n+1)} + O(\lambda^4 2^{-4n}))\, dF(\lambda)$
$\qquad\qquad + 2^{2n+1} \int_{2^{n/4}}^{\infty} (1 - \cos \lambda 2^{-n})\, dF(\lambda)$,

(iii) $\int_0^{2^{n/4}} \lambda^2\, dF(\lambda) + O(2^{-n}) + 2^{2n+1} \int_{2^{n/4}}^{\infty} (1 - \cos \lambda 2^{-n})\, dF(\lambda)$.

The last term of (iii) is positive and therefore $\int_0^\infty \lambda^2\, dF(\lambda) < \infty$. If $\int_0^\infty \lambda^2\, dF(\lambda)$ is assumed finite, write $\rho(2^{-n}) = \cos \pi\lambda_n 2^{-n}$ for $0 < \lambda_n < 2^n$ and n sufficiently large (we omit the case $\rho \equiv 1$). Then

$$1 + \rho''(\xi)2^{-(2n+1)} = 1 - \lambda_n^2 \pi^2 2^{-(2n+1)} + O(\lambda_n^4 2^{-4n}), \qquad 0 < \xi < 2^{-n},$$

so that

$$1 = -\rho''(\xi)\lambda_n^{-2}\pi^{-2} + O(\lambda_n^2 2^{-2n}).$$

Now as $n \to \infty$, $\xi \to 0$ and $\lambda_n^2 2^{-2n} \to 0$ since $\cos \pi\lambda_n 2^{-n} \to 1$. Thus

$$\lambda_n = (2^n/\pi) \text{ arc cos } \rho(2^{-n}) \to (1/\pi)(-\rho''(0))^{\frac{1}{2}}.$$

Lemmas 1, 2 and 3 together prove the

THEOREM. *Let* $\{X(t),\ t \varepsilon [0, 1]\}$ *be a real separable stationary Gaussian process having continuous sample paths, mean value function zero and covariance function* ρ, $\rho(0) = 1$. $X(\cdot)$ *has* a.s. *no tangential zeros and if* N *is the number of crossing zeros of* $X(\cdot)$,

$$\begin{aligned} EN &= (1/\pi)(-\rho''(0))^{\frac{1}{2}} && \text{if } \rho''(0) \text{ exists,} \\ &= +\infty && \text{if not.} \end{aligned}$$

REFERENCES

[1] ANDERSON, T. W. (1958). *An Introduction to Multivariate Analysis*. Wiley, New York.

[2] BULINSKAYA, E. V. (1961). On the mean number of crossings of a level by a stationary Gaussian process. *Teor. Veroyatnost. i Primenen.* **6** 435–438.

[3] DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[4] GRENANDER, Ulf and ROSENBLATT, MURRAY (1957). *Statistical Analysis of Stationary Time Series*. Wiley, New York.

[5] IVANOV, V. A. (1960). On the average number of crossings of a level by the sample functions of a stochastic process. *Teor. Veroyatnost. i Primenen.* **5** 319–323.

[6] KAC, M. (1943). On the average number of real roots of a random algebraic equation. *Bull. Amer. Math. Soc.* **49** 314–320.

[7] LEADBETTER, M. R. and CRYER, J. D. (1964). On the mean number of curve crossings by non-stationary normal processes. Tech. Rpt. No. 1 (Nonr 4565(00)) Research Triangle Institute.

[8] RICE, S. O. (1945). Mathematical analysis of random noise. *Bell System Tech. J.* **24** 46–156.