ON THE NUMBER OF SOLUTIONS OF SYSTEMS
OF RANDOM EQUATIONS

BY DAVID R. BRILLINGER

University of California, Berkeley

Let $\{f(x, \omega); x \in R^n, \omega \in \Omega\}$ be an n vector-valued stochastic process defined over a probability space $(\Omega, \mathscr{A}, \mu)$. Let $N(f \mid A, y)$ denote the number of elements in the set $A \cap f^{-1}(y)$, that is the number of distinct solutions of the system of equations $f(x, \omega) = y$ for $x, y \in R^n$. We develop expressions for $E\{N(f \mid A, y)\}$ and certain higher-order moments of $N(f \mid A, y)$ under regularity conditions.

1. Introduction. A variety of statistical properties have been developed for the number of solutions of an equation

$$(1.1) \qquad\qquad f(x) = y$$

in the case that $x, y \in R$ and f is a random function. See, for example, Kac (1943), Rice (1945), Cramér and Leadbetter (1967). Properties have also been developed in the case that $x, y \in C$ and f is a random analytic function, see Paley and Wiener (1934, page 178), Littlewood and Offord (1948), Offord (1965), Offord (1967). In this case (1.1) is equivalent with two real random equations in two real unknowns. Here we determine the expected value and the factorial moments of the number of solutions of n real random equations in n real unknowns under regularity conditions. The results obtained have application to the investigation of the number of extreme points of a random surface defined over R^n, for the extreme points are the solutions of the n equations resulting from setting the first derivatives of the surface to zero. We note that Longuett-Higgins (1957) has investigated the expected number of extreme points for certain random surfaces.

The proofs of the lemmas and theorems of Sections 2 and 3 of the paper have been collected in Section 4.

2. The non-stochastic case. In this section we develop an expression for the number of solutions of a system of n fixed equations in n unknowns. The expression provides a generalization of one due to Kac (1943). In what follows; if $y = (y_1, \cdots, y_n) \in R^n$, the region $|y_1|, \cdots, |y_n| < \varepsilon$ is denoted $|y| < \varepsilon$. If $A \subset R^n$, and f maps R^n into R^n, the restriction of f to A is denoted $f \mid A$. If $f: R^n \to R^n$ is Lipschitz, see Federer (1969), its Jacobian determinant existing almost everywhere is denoted Jf. The number of elements in the set $A \cap f^{-1}(y)$, $y \in R^n$, $A \subset R^n$, is denoted $N(f \mid A, y)$. This is the desired number of distinct solutions of (1.1) in the set A.

For $\varepsilon > 0$, $v \in R^n$ set

(2.1) $$N_\varepsilon(f \mid A, y) = (2\varepsilon)^{-n} \int_{|v| < \varepsilon} N(f \mid A, y + v) dv .$$

Also set

(2.2) $$\psi_\varepsilon(v) = 1 \qquad \text{for } |v| < \varepsilon ,$$
$$= 0 \qquad \text{otherwise} .$$

Then we have,

LEMMA 2.1. *Let A be a measurable subset of R^n and $f : A \to R^n$ be Lipschitz. Then*

(2.3) $$N_\varepsilon(f \mid A, y) = (2\varepsilon)^{-n} \int_A \psi_\varepsilon[f(x) - y] |Jf(x)| dx .$$

If in addition

(2.4) $$\int_A |Jf(x)| dx < \infty ,$$

then

(2.5) $$N(f \mid A, y) = \lim_{\varepsilon \to 0} (2\varepsilon)^{-n} \int_A \psi_\varepsilon[f(x) - y] |Jf(x)| dx$$

for almost all y and indeed

(2.6) $$\lim_{\varepsilon \to 0} \int |N(f \mid A, y) - N_\varepsilon(f \mid A, y)| dy = 0 .$$

Finally if $N(f \mid A, u)$ is continuous $u = y$, then (2.5) folds for that y.

Expression (2.5) is the promised formula for the number of solutions of interest. The next lemma indicates one set of conditions under which $N(f \mid A, u)$ is continuous at $u = y$. We say that a continuously differentiable $f : R^n \to R^n$ is *normal* above $y \in R^n$ if $Jf(x) \neq 0$ for $x \in f^{-1}(y)$, (see Whitney (1957), page 145).

LEMMA 2.2. *Let A be an open bounded subset of R^n and let $f : A \to R^n$ be normal over y. Then $N(f \mid A, u)$ is continuous at $u = y$.*

This lemma, together with Lemma 2.1, indicates that (2.5) holds for given y if f is normal above y.

3. The stochastic case. We now turn to a determination of the mean number of solutions of a random equation $f(x, \omega) = y$ falling in a set A in the case that $f(x, \omega)$ is a vector-valued stochastic process. We have,

THEOREM 3.1. *Let A be a measurable subset of R^n. Let $\{f(x, \omega); x \in R^n; \omega \in \Omega\}$ be an n vector-valued stochastic process defined over a probability space $(\Omega, \mathscr{A}, \mu)$. Let $f(x, \omega)$ be Lipschitz with probability one for $x \in A$. Let the variates $\alpha = f(x, \omega)$, $\beta = Jf(x, \omega)$ have joint density $p(\alpha; \beta; x)$, $\alpha \in R^n$, $\beta \in R$, $x \in A$, satisfying*

(3.1) $$\int \int \int_A |\beta| p(\alpha; \beta; x) d\alpha d\beta dx < \infty .$$

Then

(3.2) $$E\{N(f \mid A, y)\} = \int \int_A |\beta| p(y; \beta; x) d\beta dx$$

for almost all $y \in R^n$.

Expression (3.2) was set down by Rice (1945) in the case $n = 1$. We remark that if $p(\alpha; x)$ denotes the density of $\alpha = f(x, \omega)$, then an alternate form for (3.2), involving a conditional expected value, is

$$(3.3) \qquad E\{N(f \mid A, y)\} = \int_A p(y; x)E\{|Jf(x)| : f(x) = y\}dx \,.$$

The solutions of $f(x, \omega) = y$ determine a multidimensional point process in $R^n \cdot$ (These are discussed in Srinivasavan (1969).) If A is taken to be a small parallelipiped of volume $|A|$ and $x \in A$, then from (3.2)

$$(3.4) \qquad E\{N(f \mid A, y)\} = |A| \int |\beta| p(y; \beta; x)d\beta \,,$$

showing that the intensity parameter of this point process is $\int |\beta| p(y; \beta; x)d\beta$.

One application of (3.2) is to provide a bound for crossing probabilities of the form $\text{Prob}[f(x, \omega) = y$ for some $x \in A]$. Clearly this probability is less than or equal to $E\{N(f \mid A, y)\}$. We may conclude, for example, that the probability is zero if (3.2) holds and $p(y; \beta; x) = 0$ for almost all $\beta \in R^n$, $x \in A$.

Theorem 3.1 provides the expected number of solutions for almost all $y \in R^n$. If some particular value of y is of interest, then the following result may be of use.

COROLLARY 3.1. *Under the conditions of the theorem and if* (i) $N(f \mid A, u)$ *is continuous at* $u = y$ *with probability one,* (ii) $E\{N(f \mid A, u)^{1+\delta}\} < \infty$ *for some* $\delta > 0$ *and for u in a neighborhood of y,* (iii) $\int\int_A |\beta| p(u; \beta; x)d\beta dx$ *is continuous at* $u = y$, *then* (3.2) *holds.*

We remark that it follows from Lemma 2.2 that (i) holds if the sample paths $f(x; \omega)$ are normal over y for almost all ω.

We next turn to the investigation of a function related to the higher order moments of the number of solutions. Given measurable subsets $A_1, \cdots, A_k$ of R^n and $f: R^n \to R^n$ consider the number of solutions of the system of equations

$$(3.5) \qquad f(x_1) = y_1, \cdots, f(x_k) = y_k$$

for $y_1, \cdots, y_k \in R^n$ with $x_j \in A_j$, $x_i \neq x_j$, $1 \leq i < j \leq k$. In the case that the A_j are disjoint, the number of solutions is

$$(3.6) \qquad N(f \mid A_1, y_1) \cdots N(f \mid A_k, y_k) \,.$$

In the case that $A_j = A$, $y_j = y$, $N = N(f \mid A, y)$ the number of solutions is

$$(3.7) \qquad N(N - 1) \cdots (N + k + 1) \,.$$

Letting $B = \{(x_1, \cdots, x_k) : x_j \in A_j, x_i \neq x_j, 1 \leq i < j \leq k\}$, denoting the map of (3.5) by $\bar{f}: R^{nk} \to R^{nk}$ and letting $N(f \mid B, y_1, \cdots, y_k)$ denote the number of solutions of (3.5) falling in B we have,

THEOREM 3.2. *Let* $A_1, \cdots, A_k$ *be measurable subsets of* R^n. *Let* $\{f(x, \omega);$ $x \in \bar{R}^n$, $\omega \in \Omega\}$ *be an n vector-valued stochastic process defined over the probability space* $(\Omega, \mathscr{S}, \mu)$. *Let* $f(x, \omega)$ *be Lipschitz with probability one for* $x \in A_1 \cdots A_k$.

Let the variates $\alpha_j = f(x_j, \omega)$, $\beta_j = Jf(x_j, \omega)$, $j = 1, \cdots, k$ *have joint density* $p(\alpha_1, \cdots, \alpha_k; \beta_1, \cdots, \beta_k; x_1, \cdots, x_k)$ *for distinct* x_j, $x_j \in A_j$ *with*

$$(3.8) \quad \int \cdots \int \int_{A_1} \cdots \int_{A_k} |\beta_1| \cdots |\beta_k| p(\alpha_1, \cdots, \alpha_k; \beta_1, \cdots, \beta_k; x_1, \cdots, x_k)$$
$$\times \prod_1^k (d\alpha_j d\beta_j dx_j)$$

finite. Then

$$E\{N(\tilde{f} \mid B, y_1, \cdots, y_k)\} = \int \cdots \int \int_{A_1} \cdots \int_{A_k} |\beta_1| \cdots |\beta_k|$$
$$(3.9) \quad \times p(y_1, \cdots, y_k; \beta_1, \cdots, \beta_k; x_1, \cdots, x_k)$$
$$\times \prod_1^k (d\beta_j dx_j)$$

for almost all $y_1, \cdots, y_k \in R^n$.

As one implication of this theorem, we mention that if $A_1, \cdots, A_k$ are small disjoint parallelipipeds of volumes $|A_1|, \cdots, |A_k|$ and $x_j \in A_j$, then

$$E\{N(f \mid A_1, y_1) \cdots N(f \mid A_k, y_k)\}$$
$$(3.10) \quad = |A_1| \cdots |A_k| \int \cdots \int |\beta_1| \cdots |\beta_k|$$
$$\times p(y_1, \cdots, y_k; \beta_1, \cdots, \beta_k; x_1, \cdots, x_k) d\beta_1 \cdots d\beta_k$$

and so

$$(3.11) \quad \int \cdots \int |\beta_1| \cdots |\beta_k| p(y_1, \cdots, y_k; \beta_1, \cdots, \beta_k; x_1, \cdots, x_k) d\beta_1 \cdots d\beta_k$$

may be interpreted as a product density of order k (see Srinivasavan (1969)) of the multidimensional point process resulting from the solutions of (3.5).

If one is interested in the factorial moment of order k of $N = N(f \mid A, y)$ for some prespecified y one has,

COROLLARY 3.2. *Under the conditions of the theorem and if* (i) $N(\tilde{f} \mid B, u_1, \cdots, u_k)$ *is continuous at* $(u_1, \cdots, u_k) = (y, \cdots, y)$ *with probability one,* (ii) $E\{N(\tilde{f} \mid B, u_1, \cdots, u_k)^{1+\delta}\} < \infty$ *for some* $\delta > 0$ *and for* $u_1, \cdots, u_k$ *in a neighborhood of* $(y, \cdots, y)$, (iii) $\int \cdots \int \int_A \cdots \int_A |\beta_1| \cdots |\beta_k| p(u_1, \cdots, u_k; \beta_1, \cdots, \beta_k; x_1 \cdots, x_k) d\beta_1 \cdots d\beta_k dx_1 \cdots dx_k$ *is continuous at* $(u_1, \cdots, u_k) = (y, \cdots, y)$, *then*

$$E\{N(N - 1) \cdots (N - k + 1)\}$$
$$(3.12) \quad = \int \cdots \int \int_A \cdots \int_A |\beta_1| \cdots |\beta_k|$$
$$\times p(y, \cdots, y; \beta_1, \cdots, \beta_k; x_1, \cdots, x_k) d\beta_1 \cdots d\beta_k dx_1 \cdots dx_k .$$

We mention that $\tilde{f}$ will be normal above $(y, \cdots, y)$ when f is normal above y and so following Lemma 2.2 (i) above will hold in the case that $A_1, \cdots, A_k$ are open and bounded and f is normal above y with probability one. Expression (3.12) was set down by Cramér and Leadbetter (1967) in the case of Gaussian $f(x, \omega)$ and $n = 1$.

4. Proofs. We begin with a proof of Lemma 2.1.

PROOF OF LEMMA 2.1. Kirszbaum's Theorem (see Federer (1969), page 201)

indicates the existence of a Lipschitz extension of f with domain R^n. Theorem 3.2.5 of Federer (1969) (or Theorem 2, page 374 in Rado and Reichelderfer (1955)) then applies to give

$$(4.1) \qquad \int_A g[f(x)]|Jf(x)|dx = \int_{R^n} g(u)N(f\,|\,A, u)du$$

for measurable $g: R^n \to R$. Taking $g(u) = (2\varepsilon)^{-n}\phi_\varepsilon(u - y)$ in (4.1) gives (2.3) after a change in variable.

Taking $g(u) = 1$, shows that

$$(4.2) \qquad \int N(f\,|\,A, u)du = \int_A |Jf(x)|dx$$

and so $N(f\,|\,A, u)$ is integrable in view of (2.4). The conclusions of the lemma now follow from a standard convergence theorem (see, for example, Theorems 1.1.1, 1.3.2 in Bochner (1960).)

PROOF OF LEMMA 2.2. Under the stated conditions the set of solutions can have no limit points for the Jacobian would then vanish at some solution. The solutions are therefore isolated and finite in number. The Inverse Function Theorem then applies to give the existence of a continuously differentiable inverse in the neighborhood of each solution. If y is altered by a sufficiently small amount, it follows that the number of solutions is unchanged and so N is continuous.

PROOF OF THEOREM 3.1. We begin by noting, from (2.1), (4.2), that

$$(4.3) \qquad \int N(f\,|\,A, u)du, \quad \int N_\varepsilon(f\,|\,A, u)du = \int_A |Jf(x)|dx$$

and therefore, in view of (3.1),

$$(4.4) \qquad E\{\int N(f\,|\,A, u)du\}, \quad E\{\int N_\varepsilon(f\,|\,A, u)du\} < \infty.$$

In consequence, it follows from bounded convergence, Fubini's Theorem and (2.6) that

$$(4.5) \qquad \begin{aligned} &\lim_{\varepsilon \to 0} \int |E\{N(f\,|\,A, u)\} - E\{N_\varepsilon(f\,|\,A, u)\}|du \\ &= \lim_{\varepsilon \to 0} E\{\int|N(f\,|\,A, u) - N_\varepsilon(f\,|\,A, u)|du\} \\ &= 0 \end{aligned}$$

At the same time, we have from (2.3),

$$(4.6) \qquad E\{N_\varepsilon(f\,|\,A, y)\} = (2\varepsilon)^{-n} \int\int\int_A \phi_\varepsilon(\alpha - y)|\beta|p(\alpha; \beta; x)d\alpha d\beta dx$$

and so

$$(4.7) \qquad \lim_{\varepsilon \to 0} \int |E\{N_\varepsilon(f\,|\,A, y)\} - \int\int_A |\beta|p(y; \beta; x)d\beta dx|dy = 0$$

(by Theorem 1.3.2 of Bochner (1960).) Expressions (4.5) and (4.6) together now give

$$(4.8) \qquad \int |E\{N(f\,|\,A, y)\} - \int\int_A |\beta|p(y; \beta; x)d\beta dx|dy = 0$$

and thence the conclusion of the theorem.

PROOF OF COROLLARY 3.1. Under the stated conditions

$$E\{N(f \mid A, y)\} = E\{\lim_{u \to y} N(f \mid A, u)\}$$
$$= \lim_{u \to y} E\{N(f \mid A, u)\}$$
$$= \iint_A |\beta| p(y; \beta; x) d\beta dx.$$

PROOF OF THEOREM 3.2. B is a measurable subset of R^{nk}. The Jacobian of the map $\tilde{f}$ is given by $Jf(x_1) \cdots Jf(x_k)$. The conclusion of the theorem now follows directly from Theorem 3.1 taking n to be nk, A to be B and f to be $\tilde{f}$.

PROOF OF COROLLARY 3.2. This result follows directly from Corollary 3.1 in the above manner.

5. Concluding remarks. We note here that the results obtained are easily modified, in the manner of Leadbetter (1966), to yield the moments of the number of solutions of the equation

(5.1) $$f(x) = g(x)$$

for a fixed measurable n vector-valued function g.

The reader will have noted that the results obtained required expression (4.1) in an essential manner. In fact Federer, Theorem 3.2.5 develops expression (4.1) in the more general setting of maps $f: R^m \to R^n$ with $m \leqq n$ using Hausdorf m-measure. This suggests the possibility of extending the Theorems of this paper to apply to n vector-valued stochastic processes $f(x, \omega)$, $x \in R^m$, $m \leqq n$.

In another direction we mention that if A is a bounded open set, $f: \bar{A} \to R^n$ is continuously differentiable and $\mu(f \mid A, y)$ is the topological index of the mapping f with domain A at the point y (see Rado and Reichelderfer (1955), page 125), then as an analog of (4.1) one has

(5.2) $$\int_A g[f(x)]Jf(x)dx = \int g(u)\mu(f \mid A, u)du$$

(*ibid.* page 374) and so one has, for example, under the conditions of Theorem 3.1

(5.3) $$E\{\mu(f \mid A, y)\} = \iint_A \beta p(y; \beta; x) d\beta dx.$$

I would like to thank Professor M. W. Hirsch for suggesting Lemma 2.2 to me.

REFERENCES

[1] BOCHNER, S. (1960). *Harmonic Analysis and the Theory of Probability.* Univ. of California Press.
[2] CRAMÉR, H. and LEADBETTER, M. R. (1967). *Stationary and Related Stochastic Processes.* Wiley, New York.
[3] FEDERER, H. (1969). *Geometric Measure Theory.* Springer-Verlag, Berlin.

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

[5] LEADBETTER, M. R. (1966). On crossings of levels and curves by a wide class of stochastic processes. *Ann. Math. Statist.* **37** 260–267.

[6] LITTLEWOOD, J. E. and OFFORD, A. C. (1948). On the distribution of zeros and *a*-values of a random integral function II. *Ann. of Math.* **49** 885–952.

[7] LONGUET-HIGGINS, M. S. (1957). The statistical analysis of a random moving surface. *Philos. Trans. Roy. Soc. London Ser. A* **249** 321–387.

[8] OFFORD, A. C. (1965). The distribution of the values of an entire function whose coefficients are independent random variables I. *Proc. London Math. Soc.* (3) **14A** 199–238.

[9] OFFORD, A. C. (1967). The distribution of zeros of power series whose coefficients are independent random variables. *Indian J. Math.* **9** 175–196.

[10] PALEY, R. E. A. C. and WIENER, N. (1934). *Fourier Transforms in the Complex Domain.* Amer. Math. Soc, Providence.

[11] RADO, T. and REICHELDERFER, P. V. (1955). *Continuous Transformations in Analysis.* Springer-Verlag, Berlin.

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

[13] SRINIVASAN, S. K. (1969). *Stochastic Theory and Cascade Processes.* Elsevier, New York.

[14] WHITNEY, H. (1957). *Geometric Integration Theory.* Princeton Univ. Press.

[Note added in proof. Beljaev (1967), *Soviet Math. Dokl.* **8** 1107–1109, has announced results that are extensions of those of Longuet-Higgins (1957).]