ON THE PATHOLOGICAL CHARACTER OF INDEPENDENT RANDOM VARIABLES

by

DONALD GEORGE BRENNAN

B.S., M.I.T.

(1955)

SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF

PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF

TECHNOLOGY

May, 1959

Signature of Author . . . .
D. G. Brennan
Department of Mathematics, May 15, 1959

Certified by. . . . .
Norbert Wiener
Thesis Supervisor

Accepted by . . . .
W. T. Martin
Chairman, Departmental Committee on Graduate Students

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE 39, MASS.

DEPARTMENT OF MATHEMATICS

April 22, 1959

Mr. Donald G. Brennan
Department of Mathematics
M.I.T.
Cambridge 39, Massachusetts

Dear Mr. Brennan:

This letter is to give you permission to print additional copies of your thesis by the Multilith process, and to submit to the Department of Mathematics copies thus produced, in lieu of the typed copies normally required.

A copy of this letter is to be reproduced by the same process and is to be placed in each copy of the thesis immediately following its title page.

Sincerely yours,

G. B. Thomas, Jr.
for the
Department Graduate Committee

GBT:pr

ON THE PATHOLOGICAL CHARACTER

OF INDEPENDENT RANDOM VARIABLES

D. G. Brennan

Submitted to the Department of Mathematics on May 15, 1959, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

The following theorem is proved: If f and g are R_n-valued random variables on an arbitrary probability space, and if f and g are measure-continuous (i.e., send sets of small measure into sets of small measure) and have image sets in R_n of positive outer measure, then f and g are not independent. This theorem, which appears to be the first result of this type obtained since random variables were first adequately defined, provides considerable information about various probability-theoretic questions related to independence or the lack thereof. Some of this information is exhibited in a sequence of remarks, including: (a) The theorem is at least close to being "best possible". (b) Independent r.v.'s on product spaces are independent essentially because they collapse the dimension of the domain. (c) Two continuous and measure-continuous real r.v.'s on a connected topological probability space are independent if and only if at least one is constant. (d) In particular, absolutely continuous r.v.'s on the unit interval are continuous and measure-continuous. (e) The theorem has important consequences for applications of probability theory to physics and engineering.

Thesis supervisor: Norbert Wiener, Institute Professor

I. INTRODUCTION

With respect to the usual axiomatic framework of modern probability theory, in which random variables are taken to be measurable functions on a totally finite measure space, it is fairly well known that the structure of the measure space underlying any particular development is by no means completely arbitrary. See, for example, the discussion by J. L. Doob in Appendix I of [1]. But it is less widely appreciated that the random variables themselves, occurring in such developments, are also subject to very stringent constraints, especially in the case of independent random variables. There appears to be no information in the literature about the function-theoretic properties of independent random variables, even when the underlying space is quite simple, such as an interval or product of intervals under Lebesque measure.

The objective of this paper is to set forth a theorem concerning the function-theoretic character of independent random variables defined on a completely arbitrary probability space. This theorem, which seems to be the first result obtained in this direction since random variables were adequately defined about 30 years ago, provides a good deal of illumination on various probability-theoretic questions related to independence or the lack thereof. Several of these questions will be exhibited. In special cases, the theorem provides a complete characterization of independence. It will be seen that the structure of independent random variables is such as could reasonably be described as "pathological" from the point of view of ordinary function theory; e.g., real-valued random variables on the unit interval I that are C^1 on I cannot be independent unless constant, and there are no non-trivial, complex-valued independent random variables on the unit square $I \times I$ that are analytic on $I \times I$.

II. DEFINITIONS AND STATEMENT OF THE THEOREM

We recall that a probability space (Ω, A, P) is a collection of a point set Ω, a sigma-ring A of subsets of Ω such that $\Omega \in A$, and a measure P defined on A such that $P(\Omega) = 1$. A real-valued random variable on (Ω, A, P) is a function f from Ω to R_1 that is measurable with respect to A, i.e., if $E \subseteq R_1$ is measurable, then $f^{-1}(E) \in A$. More generally, R_n-valued random variables are functions from Ω to R_n satisfying the same condition. (It is immaterial for our purposes whether the measurable sets of R_n are taken as the Lebesgue-measurable sets or only the Borel sets, or whether the measure space (Ω, A, P) is completed or not.) We say that two random variables f and g are independent if $P[f^{-1}(E_1) \cap g^{-1}(E_2)] = P[f^{-1}(E_1)]P[g^{-1}(E_2)]$ whenever E_1 and E_2 are measurable sets in R_n. Although random variables are point functions, in writing $f(E)$ we shall as usual mean the set $\left\{f(\omega) | \omega \in E\right\}$ in R_n, whenever E is a set in Ω. We shall write μ for n-dimensional Lebesgue measure and $\overline{\mu}$ for outer Lebesgue measure. We shall usually write Ω for a probability space, with the understanding that A and P are implied.

One of the key ideas in the development below is that of "measure-continuity", meaning the property of sending sets of small measure into sets of small measure, more precisely formulated in the

<u>Definition</u>. Let $f: \Omega \to R_n$ be a random variable. We say that f is <u>measure-continuous</u> if, whenever $\epsilon > 0$ is given, there is a $\delta > 0$ such that $P(E) < \delta$ implies $\overline{\mu}\{f(E)\} < \epsilon$. We say that f is <u>weakly measure-continuous</u> if f sends no sets of zero P-measure into sets of positive

outer μ-measure.

Of course, measure-continuity can be formulated for functions (which need not be measurable) between any two measure spaces; however, the setting employed in this definition will suffice for present purposes. But even a measurable function need not send measurable sets in its domain into measurable sets in its range, which accounts for the presence of the outer measure in the definition. It is obvious that any measure-continuous f is weakly measure-continuous, but the converse is false. It should be noted that measure-continuity neither implies nor is implied by absolute continuity (with respect to μ) of the measure induced in R_n by f. Also, in cases where the measure space Ω is equipped with a topology, there is in general no relation between topological continuity and measure-continuity of f; for example, a real-valued measure-continuous f on the unit interval need not be continuous at any point, and the Cantor function [2] shows that even a monotone continuous function need not be weakly measure-continuous.

We are now ready to state our main result.

<u>Theorem</u>. Let (Ω, A, P) be an arbitrary probability space and let f and g be measure-continuous random variables on Ω with $\overline{\mu}\{f(\Omega)\} > 0$, $\overline{\mu}\{g(\Omega)\} > 0$. Then f and g are not independent.

What we shall actually prove is the slightly sharper but less symmetric

<u>Theorem'</u>. Let (Ω, A, P) be a probability space and let f and g be random variables on Ω with $\overline{\mu}\{f(\Omega)\} > 0$, $\overline{\mu}\{g(\Omega)\} > 0$. If f is weakly measure-continuous and g is measure-continuous, then f and g are not independent.

(In speaking of "the theorem" without further qualification, either the symmetric form or the sharper form may be substituted.)

Thus, independent random variables cannot have "thick" image sets if they are required to fill up their image sets "smoothly". Before presenting the proof of this theorem, we shall provide a sequence of observations and applications to illustrate some of the content of the theorem.

III. VARIOUS REMARKS AND CONSEQUENCES

Remark 1. There is no deep significance attached to the use of (R_n, Lebesgue) for the range space. As can be seen from the proof, the only vital property required of the range space is that it be a measure space that admits a systematic way of splitting it up into disjoint pieces of arbitrarily small measure, and this will be the case under very general conditions. The measure-continuity of the random variables would then be formulated in terms of the given measure (and its induced outer measure) on the range space. We shall not pursue this here, however, and (R_n, Lebesgue) will be used throughout as the range space.

Remark 2. Although the theorem is true in the generality stated above, it is interesting to note that the theorem has no content in cases where Ω is a finite or countable set. There is thus some sense in which the notion of independence is most "natural" for discrete random variables.

Remark 3. If the hypotheses of the theorem are regarded as two separate conditions each on f and g, then no one of the four conditions can be dropped. This can be proved with simple real random variables ($n = 1$) with Ω taken to be the unit interval $I = [0, 1]$ with Lebesgue measure, as follows. i) Let f be the trigonometric function $f(t) = \sin 2\pi t$, $t \in I$, and let g be the second Rademacher function $g(t) = \operatorname{sgn}(\sin 4\pi t)$. Then it is easy to see (draw pictures) that a) f and g are independent, b) f and g are measure-continuous, c) $\mu\{f(\Omega)\} = 2 > 0$, d) $\mu\{g(\Omega)\} = 0$. (In this example, both r.v.'s have non-degenerate distributions; an even simpler example with $\mu\{g(\Omega)\} = 0$ can be obtained by taking g = constant, as g would then be independent of any other r.v. whatever on the same space, but this g would have a degenerate distribution.) ii) Let $f(t) = \sin 2\pi t$ as before. Let $C(t)$ be the Cantor function on I and let $g(t) = C(t)$ for every t in the Cantor set and let $g(t) = 0$ on every interval in the complement of the Cantor set. Then a) f and g are independent (since g = constant a.e.), b) $\mu\{f(\Omega)\} = 2 > 0$, $\mu\{g(\Omega)\} = 1 > 0$, c) f is measure-continuous, but d) g is not weakly measure-continuous. These examples show that the theorem certainly cannot be substantially improved. Obviously the condition of positive outer measure of the image sets cannot be weakened at all; but whether or not the measure-continuity conditions can be significantly relaxed is a slightly open question, which will be further discussed below.

<u>Remark 4</u>. The usual method of constructing independent r.v.'s on product spaces leads to r.v.'s that are not measure-continuous. For example, let Ω be the unit square $I \times I$ under Lebesgue measure, let $n = 1$, and let $f(x, y) = x$, $g(x, y) = y$. Then it is well known (and also easy to verify) that f and g are independent. But if $y_o \in I$ and $x_o \in I$, then f and g respectively send the lines $I \times y_o$ and

$x_o \times I$ (of P-measure 0) onto I (of μ-measure 1), and similarly if $f(x, y) = f_o(x)$, $g(x, y) = g_o(y)$. Therefore, such (real) variables can be measure-continuous only if their complete image sets are of measure zero in the first place. Random variables on a product space that take values in a range space of the same dimension as the domain cannot be independent if they are "smooth". For example, if f and g are non-constant complex-valued random variables on $I \times I$ that are analytic on $I \times I$, then they are measure-continuous (because the Jacobians of the transformations given by f and g are respectively $|f'(z)|^2$ and $|g'(z)|^2$, which are bounded on $I \times I$) and have image sets of positive outer measure in R_2 (because non-constant analytic functions are open maps). It follows from the main theorem that such f and g cannot be independent.

Remark 5. It is not possible to formulate a theorem with the generality of the main theorem yielding "if and only if" conditions for independence in terms of measure-continuity and measures of image sets. This can be seen by considering two independent, non-degenerate r.v.'s on any Ω; by composing one of these r.v.'s with a suitable measure-preserving transformation of Ω, it is always possible to destroy the independence without affecting the measure-continuity or measure of the image sets of these variables. However, if there are additional conditions on the structure of the domain and range spaces, then it is possible to formulate "if and only if" theorems in terms of the function-theoretic structure of the r.v.'s, as in the following

Corollary. Let Ω be a connected topological probability space, and let f and g be real-valued, continuous and measure-continuous r.v.'s on Ω. Then f and g are independent if and only if at least one of f or g is constant.

Proof of Corollary: The continuous image of a connected set is connected, and the only connected sets in the line are (possibly degenerate) intervals. If both the image intervals of f and g are non-degenerate, they would both have positive measure, which completes the "only if" part; the "if" part is the same trivial fact mentioned previously that a constant r.v. is independent of any other r.v. on the same space.

Remark 6. One special case of the preceding corollary is of sufficient interest to warrant discussion on its own. This is the case where Ω is the unit interval I under Lebesgue measure, and f and g are continuous relative to the usual topology on I. Let us recall that a real function f of a real variable is said to be absolutely continuous if, whenever $\epsilon > 0$ is given, there is a $\delta > 0$ such that for every finite collection of disjoint half-open intervals $(a_1, b_1], (a_2, b_2], \ldots, (a_n, b_n]$ in the domain of f with $\Sigma \mid b_k - a_k \mid < \delta$, it follows that $\Sigma \mid f(b_k) - f(a_k) \mid < \epsilon$. It is then obvious that any absolutely continuous r.v. on I is continuous and measure-continuous, which establishes the

Sub-Corollary. Let f and g be absolutely continuous (real) random variables on I. Then f and g are independent if and only if at least one of f or g is constant.

Although any absolutely continuous function is continuous and measure-continuous, the converse does not hold. This is illustrated in Fig. 1. The (continuous) function f whose graph is given there is defined in a sequence of trapezoidal

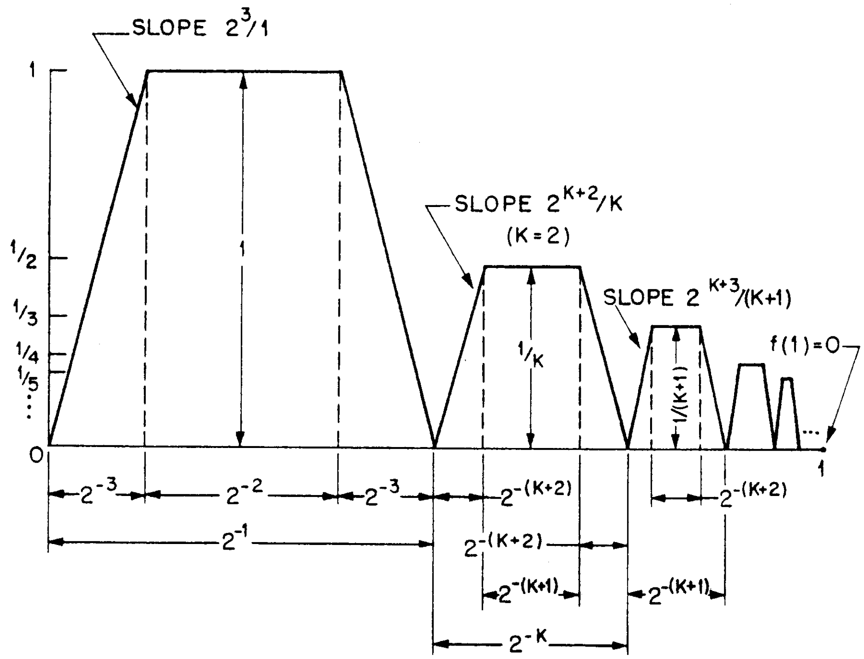

Fig. 1: Example of a continuous and measure-continuous function that is not absolutely continuous.

pieces whose base widths (2^{-k}) decrease much more rapidly than their heights ($1/k$). The slope on the sides of the $k^{\underline{th}}$ trapezoid is $2^{k+2}/k$; by considering a sequence of intervals under these sides, $k = 1, 2, 3, \ldots$, it is easy to show that f is not absolutely continuous on $[0, 1]$. However, for $\eta > 0$, f is absolutely continuous on $[0, 1 - \eta]$ and uniformly small on $[1 - \eta, 1]$; from this, it can be seen that f is measure-continuous. The essential reason why this example "works" is that the pieces $|f(b_k) - f(a_k)|$ that cause the failure of absolute continuity "overlap" when seen in the image set. It is easy to see that a monotone function that is continuous and measure-continuous is also absolutely continuous.

Hence, the "sub-corollary" above is not actually as good a theorem as results simply from specializing Ω to I in the corollary. But it is of interest to note the connection with the familiar notion of absolute continuity.

Remark 7. An attempt to construct relatively simple non-constant, independent, continuous (but not measure-continuous) real r.v.'s on I is likely to lead to the suspicion that the word "absolutely" could be deleted from the sub-corollary above. However, this is not the case. If f and g are the x and y component functions of the square-filling Peano curve [8], then f and g are (a) continuous, (b) non-constant, and (c) independent. I am indebted for this example to J. D. Lordan and Norbert Wiener, each of whom independently suggested the Peano curve. The independence of these functions can be proved by routine but long-winded computations which we omit.

Remark 8. The definition of measure-continuity used in connection with the main theorem is a "uniform" measure-continuity requirement. A natural way of formulating "local" measure-continuity for topological measure spaces would be: If $\epsilon > 0$ is given and $x \in \Omega$, then there is a neighborhood U of x and a $\delta > 0$ such that if $E \subseteq U$ and $P(E) < \delta$, then $\overline{\mu}\{f(E)\} < \epsilon$. But any continuous function has this property, and Remark 7 shows that this is not a possible direction of sharpening the main theorem even in cases where the measure space is equipped with the Haar measure of a topological group. Whether there is some other direction of "local" measure-continuity that would lead to fruitful results is not entirely clear, but it does not seem very likely that there is much to be done in this direction.

However, there are certainly other directions of formulating similar but non-overlapping theorems, especially for the case of real random variables on I. For example, a simplified version of the proof below will easily establish: If f is strictly monotone and g is measure-continuous and $\overline{\mu}\{g(I)\} > 0$, then f and g are not independent. Also, it does not seem likely that two non-constant, continuous functions of bounded variation on I could be independent. Many variations of this type are possible.

Remark 9. In the light of the comments about r.v.'s on product spaces in Remark 4 and the behavior of the Peano functions in Remark 7, it is possible to summarize the content of the main theorem rather vaguely with the statement: Independent random variables must either "collapse" the "dimension" of their domain or else be wildly pathological. The "collapsing" (which is itself a form of pathology) can take place in various ways; e.g., in the honest sense of sending a product of one-dimensional spaces into Euclidean space of lower dimension than the product (the usual case of independent r.v.'s on product spaces), or by sending the entire domain into a set of measure zero.

Remark 10. As illustrative applications to random variables on I of prescribed functional form studied elsewhere in the literature, we mention: (a) The Rademacher functions, which are well known to be independent, are obviously measure-continuous, but they all have image sets of measure zero. In consequence of the main theorem, any attempt to construct independent r.v.'s similar to the Rademacher functions but with "tilted" tops (and/or bottoms) would be doomed to failure. (b) Kac [3, 4] and Fine [5, 6] have studied asymptotic distributions of sums of "sawtooth" functions such as $2^n t - [2^n t] - 1/2$, noted by these authors to have "a certain degree of statistical dependence". The existence of such dependence is an immediate consequence of the theorem; for every n, these functions are obviously measure-continuous and have image sets of positive measure.

Remark 11. Segal [7] has criticised the usual measure-theoretic setting of random variables as follows: "Random variables are usually defined as measurable functions on probability spaces. This is a long way from either the practical statistical or intuitive conceptual formulation of the notion, especially as the probability spaces required for dealing with simple concrete situations may be mathematically relatively sophisticated." Although Segal's criticism may possibly have some merit in the case of discrete r.v.'s, it is not well taken in the case of r.v.'s intended to represent quantities that vary continuously in time. I have discussed elsewhere [9] a rather large class of interesting applications in which the underlying probability space enters in a completely natural way, and the r.v.'s on this space are as close as could be imagined to the "practical statistical" and "intuitive conceptual" notions involved. In order to demonstrate the relevance of the main theorem to such applications, and show incidentally in the process that Segal's criticism is misplaced, let us consider a representative application of this type. We must introduce some physical considerations for this purpose.

The output voltage of a random noise generator is often said to have a "distribution" of specified form, say the normal distribution. So far as such a statement ever has any operational significance, it means the following: If the output voltage of the generator is recorded on a chart recording or oscillogram throughout an interval of suitable length in time, and the waveform is "sampled" at closely spaced intervals, and the fraction of sampled values $\leq x$ is counted, the resulting (step-) function of x will be closely approximated by the normal

(Gaussian) distribution function. It is easy to see that this "sampled" distribution function converges (weakly) to the distribution function of that random variable on I whose graph is identical to the given chart recording. It follows that any physically measurable statistical properties of the noise generator will be precisely properties of real r.v.'s on I, the said r.v.'s having whatever continuity or other analytical properties are appropriate to the output voltage of the generator. Exactly the same situation prevails in similar applications in meteorology, communication, oceanography, seismology, ionosphere physics, the study of electroencephalograms, and more generally wherever the statistical study of data given continuously in time is of interest. We believe that this is sufficient in itself to place Segal's remarks in their proper light.

However, there is more to be said. Consider the "shot" noise present in the output circuit of a vacuum tube, consisting of the sum of pulses of current due to the arrival of individual electrons at the collector anode of the tube. In most tubes, there would be between 10^{15} and 10^{20} electrons per second arriving at the anode, and, the individual pulses of current being "small" and "random", it is universally argued in physics and engineering texts and papers on the subject that the sum of these pulses--i.e., the total output current--must have a Gaussian distribution, in consequence of the central limit theorem. But let us look more closely at both the physics and the mathematics of this situation. The operational significance of the "Gaussian distribution" is the same as before, i.e., it means that sections in time of the output current considered qua random variables over intervals with Lebesgue measure should have approximately Gaussian distributions. The individual current pulses constituting the summands are also r.v.'s relative to such intervals, and the total output current is simply the pointwise sum of such functions. Now, the magnitude of each current pulse as the corresponding electron nears the anode is proportional to the velocity of the electron, so the time derivative of the current pulse is proportional to the acceleration of the electron. This acceleration is evidently bounded, since the mass of the electron is bounded below by the rest mass of the electron and the available forces acting on the electron are bounded--outside the anode by the accelerating potential of the tube, inside the anode by the gradient of the potential barrier at the anode surface. It follows that the appropriate mathematical functions to represent the current pulses would be continuous functions with first derivatives that were bounded uniformly in time and over all electrons. But every such function is absolutely continuous, and these same functions are the r.v.'s constituting the summands of the random variable representing the total output current. It follows from the sub-corollary to the main theorem that the summand r.v.'s cannot be independent. Therefore, all of the standard forms of the central limit theorem are inapplicable to this case.

A few central-limit type results have been obtained in special cases where independence of all summands was not assumed. Most of these assume independence "sufficiently far out", i.e., they assume a sequence of r.v.'s such that a given r.v. in the sequence is independent of all the r.v.'s in the sequence beyond some later point in the index set. It is clear that this setting is of no use in the problem above. Kac [3, 4] and Fine [5, 6] have studied asymptotic distributions of sums of the "sawtooth" functions mentioned above and related functions of the form $f(2^n t)$, where f in the first instance is periodic of period one on the line and the $g_n(t) = f(2^n t)$ are considered as r.v.'s on I. These are not actually independent "sufficiently far out", but they "approach" independence for large n; it is easy to see that they are not measure-continuous uniformly in the index n. In contrast, the uniform bound on the derivatives of the r.v.'s representing the current pulses implies that a fixed δ in the measure-continuity defintion would suffice for all the r.v.'s of the family. Presumably one should call a family of r.v.'s satisfying this condition "equi-measure-continuous". The problem then is that there appear to be no known theorems giving results on the asymptotic distribution of sums of equi-measure-continuous r.v.'s, whereas the "current pulse" r.v.'s <u>are</u> equi-measure-continuous, and have image sets of positive measure. Moreover, this "shot noise" example is by no means isolated; it appears that most or all theoretical applications in physics and engineering where "the central limit theorem" is invoked for sums of continuously varying quantities suffer from this same problem. It is not at all impossible that this fact is closely related to the fact that many experimentalists in recent years have obtained measured distributions of noise-like quantities, and found them distinctly non-Gaussian. On the other hand, some noise-like quantities <u>have</u> been found to have approximately Gaussian distributions, which offers some hope of positive results on the unsolved

<u>Central Limit Problem</u>. To what extent, if any, can one characterize

the conditions under which normalized sums of equi-measure-continuous r.v.'s with image sets of positive outer measure have asymptotic distributions of prescribed form, especially Gaussian? In particular, what can be said for such r.v.'s that are absolutely continuous on I?

Let us note that the entire development of this paper is a direct consequence of taking seriously the idea that a random variable is a measurable function on a probability space. Simply by exploiting this fact, we have obtained interesting purely mathematical relationships between probability-theoretic properties and other mathematical properties such as absolute continuity. In turn, these relationships have indicated the existence of unsolved problems, an adequate solution to which would (a) be of intrinsic mathematical interest, and (b) have considerable importance for problems in physics and engineering. In this sense, the wheel has turned full circle: Experimental observations on dice-throwing and the raising of crops led ultimately to the formulation of an axiomatic mathematical framework for probability theory, and the apparently abstract formulation can now be seen to lead directly to results of much relevance for physics. Perhaps Segal should rather have criticised unquestioning acceptance of the idea of independence.

IV. PROOF OF THE THEOREM

The basic idea of the following proof is quite easy to see in the case of real, linear r.v.'s on I. Let, say, $f(t) = t$, $g(t) = t + 2$; by sketching the graphs of these functions, it is easy to see that they are not independent. The reasons why these particular r.v.'s are not independent extend to very general circumstances, with the aid of the following machinery. We use the notation in the statement of the theorem'.

Partition R_n into disjoint half-open cubes Q_k^m such that each cube is of length 2^{-m} per side, where k enumerates the cubes of the partition. We do this for m = 1, 2, 3, ..., where the cubes of the $(m+1)^{\underline{st}}$ partition are obtained from the $m^{\underline{th}}$ by splitting each cube into 2^n identical pieces. (E.g., for n = 1, take $Q_k^m = \left(\frac{k-1}{2^m}, \frac{k}{2^m}\right]$.) Thus $Q_k^m = \bigcup_{k'} Q_{k'}^{m+1}$, where k' runs over 2^n indices, $R_n = \bigcup_{\text{all } k} Q_k^m$ for each m, and $\mu(Q_k^m) = 2^{-mn}$ for all k. Put $E_k^m = f^{-1}(Q_k^m) \subseteq \Omega$. For each m, the E_k^m are disjoint. Let N(m) = number of indices k for which $P(E_k^m) > 0$. Applying f^{-1} to $Q_k^m = \bigcup_{k'} Q_{k'}^{m+1}$, we have $E_k^m = \bigcup_{k'} E_{k'}^{m+1}$, from which $N(m+1) \geqslant N(m)$. The weak measure-continuity of f implies that at most N(m) of the $f(E_k^m)$ have $\bar{\mu}\{f(E_k^m)\} > 0$. On the other hand, $f(E_k^m) \subseteq Q_k^m$, so $\bar{\mu}\{f(E_k^m)\} \leqslant 2^{-mn}$, and we have $\bar{\mu}\{f(\Omega)\} = \bar{\mu}\{f[\bigcup_k E_k^m]\} = \bar{\mu}\{\bigcup_k f(E_k^m)\} \leqslant \Sigma_k \bar{\mu}\{f(E_k^m)\} \leqslant N(m)2^{-mn}$, or $N(m) \geqslant 2^{mn} \bar{\mu}\{f(\Omega)\}$. Since $\bar{\mu}\{f(\Omega)\} > 0$, we have $N(m) \to \infty$ or $N(m) = \infty$ from some index m_o onwards. In the former case, since $P(\Omega) = 1$ and the E_k^m are disjoint, we see that among the N(m) E_k^m with $P(E_k^m) > 0$, there must be at least one with $0 < P(E_k^m) \leqslant 1/N(m)$. Hence, given any $\eta > 0$, there exists an m and k such that $0 < P[f^{-1}(Q_k^m)] < \eta$. The same conclusion obviously holds if $N(m) = \infty$ for $m \geqslant m_o$.

Next, let $A = g(\Omega)$. Since $\bar{\mu}(A) > 0$, the measure-continuity of g implies that there exists $\delta > 0$ such that $P(E) < \delta$ implies $\bar{\mu}\{g(E)\} < \bar{\mu}(A)/3$. (If $\bar{\mu}(A) = \infty$, choose δ so that $\bar{\mu}\{g(E)\} < 1$; it will be seen that we may henceforth assume $\bar{\mu}(A) < \infty$.) Fix δ. As above, choose m and k so that $0 < P[f^{-1}(Q_k^m)] < \delta$; fix this m and k, and put $Q = Q_k^m$, $E = f^{-1}(Q)$. Thus $\bar{\mu}\{g(E)\} < \bar{\mu}(A)/3$. Put $B = g(E)$.

There exists a measurable $B_o \supseteq B$ such that $\mu(B_o) \leq \bar{\mu}(B) + \bar{\mu}(A)/3 < 2\bar{\mu}(A)/3$. Take a measurable $A_o \supseteq A$ such that $A_o \supseteq B_o$. Let $V \equiv g(g^{-1}(A_o - B_o))$; then $V = A - B_o$. For, if $x \epsilon V$, then $x \epsilon A$, and $x = g(\omega)$ for some $\omega \epsilon g^{-1}(A_o - B_o) = g^{-1}(A_o) - g^{-1}(B_o)$, so $\omega \notin g^{-1}(B_o)$, hence $x = g(\omega) \notin B_o$, and $x \epsilon A - B_o$. If $x \epsilon A - B_o$, then $x = g(\omega)$ for some $\omega \epsilon g^{-1}(A_o)$, $\omega \notin g^{-1}(B_o)$, hence $x \epsilon V$. Thus $V = A - B_o$. Hence $A \subseteq V \cup B_o$, so $\bar{\mu}(A) \leq \bar{\mu}(V) + \mu(B_o) < \bar{\mu}(V) + 2\bar{\mu}(A)/3$, or $\bar{\mu}(V) > \bar{\mu}(A)(1 - 2/3) = \bar{\mu}(A)/3$. But $\bar{\mu}(A) > 0$, so $\bar{\mu}(V) > 0$. Hence, for $C \equiv A_o - B_o$, C is measurable, and $P[g^{-1}(C)] > 0$ since $\bar{\mu}[g(g^{-1}(C))] = \bar{\mu}(V) > 0$ and g is measure-continuous.

Finally, $C = A_o - B_o$ implies $C \cap B_o = \phi$, which implies $C \cap B = \phi$, which implies $\phi = g^{-1}(C) \cap g^{-1}(B) = g^{-1}(C) \cap E = g^{-1}(C) \cap f^{-1}(Q) = \phi$, since $E \subseteq g^{-1}(B)$. Thus $P[g^{-1}(C) \cap f^{-1}(Q)] = 0$. But $P[g^{-1}(C)] \times P[f^{-1}(Q)] > 0$, and the proof is complete.

Acknowledgments

I am indebted to Norbert Wiener for his continual willingness to provide advice and encouragement. I should also like to thank J. R. Kinney, J. J. Levin, J. D. Lordan, H. P. McKean, and Gustave Solomon for their patience in listening to preliminary expositions of portions of the above material and for their encouraging remarks.

I am also indebted to the M.I.T. Lincoln Laboratory, especially in the persons of W. B. Davenport, Jr., and W. H. Radford, for their generous financial support in the 1958-1959 academic year, during which this investigation was principally carried out.

Bibliography

[1] B. V. Gnedenko and A. N. Kolmogorov, "Limit Distributions for Sums of Independent Random Variables", Addison-Wesley, Cambridge, Mass., 1954.

[2] M. E. Munroe, "Introduction to Measure and Integration", Addison-Wesley, Cambridge, Mass., 1953; pp. 193-194.

[3] M. Kac, "Sur les fonctions $2^n t - [2^n t] - 1/2$", J. London Math. Soc., vol. 13, pp. 131-134; 1938.

[4] M. Kac, "On the distribution of values of sums of the type $\Sigma f(2^k t)$", Ann. of Math., vol. 47, pp. 33-49; 1946.

[5] N. J. Fine, "On the Walsh functions", Trans. Amer. Math. Soc., vol. 65, pp. 372-414; 1949.

[6] N. J. Fine, "On the asymptotic distribution of certain sums", Proc. Amer. Math. Soc., vol. 5, pp. 243 - 252; 1954.

[7] I. E. Segal, "Abstract probability spaces and a theorem of Kolmogoroff", Amer. Jour. of Math., vol. 76, pp. 721-732; 1954.

[8] E. W. Hobson, "The Theory of Functions of a Real Variable", vol. I, 3rd Edition, Harren Press, Washington, D. C.; p. 452 ff.

[9] D. G. Brennan, "A new approach to certain types of random functions", Physical Society Symposium on Fluctuation Phenomena and Stochastic Processes, University of London, March, 1959.

Biographical Note

D. G. Brennan was born in Waterbury, Conn., on April 9, 1926. He attended Crosby High School in Waterbury from 1938 to 1942. He entered the Massachusetts Institute of Technology as a freshman in 1951, received a B. S. in mathematics in June 1955, and completed studies for a Ph. D. in mathematics in May, 1959.

From 1942 to 1944, he was engaged in broadcast and communication system engineering, and from 1944 to 1947, he was with the U. S. Army Signal Corps, principally as Chief of Communications, Southern Islands Command, Southwest Pacific Theater. From 1947 to 1949, he was Chief Engineer of Radio Station WWCO, Waterbury, Conn. In 1949-1951, he was an instructor at the Ward School of Electronics of the University of Hartford. He has been associated with the M.I.T. Lincoln Laboratory since June 1953, where he has conducted ionosphere physics research, electronic circuit and system design, and mathematical research.

Mr. Brennan is also active in political affairs relating to science and technology. He is the Secretary of the Operating Committee on the Technical Problems of Arms Limitation of the American Academy of Arts and Sciences, has been the Branch Secretary of the Greater Boston Branch of the Federation of American Scientists since June 1957, and was recently elected to the National Council of the FAS.

He is a Registered Professional Engineer in Connecticut, a Senior Member of the Institute of Radio Engineers, and a member of Sigma Xi, the American Mathematical Society, and the Institute of Mathematical Statistics. He was a Gerard Swope Fellow in the 1955-56 academic year, and won the Boit Essay Prize at M.I.T. in 1952, 1953 and 1954.

Publications:

"On the maximum signal-to-noise ratio realizable from several noisy signals", Proc. Institute of Radio Engineers, vol. 43, p. 1530; October, 1955

"Smooth random functions need not have smooth correlation functions", Proc. IRE, vol. 45, p. 1016; July, 1957

"Phase and Amplitude Variability in Medium-Frequency Ionospheric Transmission" (with M. Lindeman Phillips), MIT Lincoln Laboratory Technical Report No. 93, September, 1957. (This is essentially a book.)

"The detection of high-altitude missile tests", pp. 171-184, in S. Melman, Editor, "Inspection for Disarmament", Columbia University Press, New York, 1958

(continued on next page)

"The extrapolation and interpolation of spatial correlation functions", International Symposium on Statistical Methods in Radio Wave Propagation, University of California at Los Angeles, June, 1958. (To be published by Pergamon Press.)

"The Detection of High-Altitude Missile Tests, II", Report of the American Academy of Arts and Sciences, Brookline, Mass., September, 1958

"A new approach to certain types of random functions", Physical Society Symposium on Fluctuation Phenomena and Stochastic Processes, University of London, March, 1959

"Outer space control and limitation of missiles", Bulletin of the Atomic Scientists, vol. 15, no. 5; May, 1959

"Linear diversity combining techniques", Proc. IRE, vol. 47, no. 6; June, 1959