ARL 66-0231
NOVEMBER 1966

Aerospace Research Laboratories

SOME TOPICS RELATED TO TRANSFORMATIONS, DISTRIBUTION FUNCTIONS AND STOCHASTIC PROCESSES

NORMAN C. SEVERO
PAUL J. SCHILLO
ROBERT H. RODINE
STATE UNIVERSITY OF NEW YORK AT BUFFALO
BUFFALO, NEW YORK

Contract No. AF 33(657)-9885
Project No. 7071

Distribution of this document is unlimited

OFFICE OF AEROSPACE RESEARCH
United States Air Force

[redacted]

ARL 66-0231

SOME TOPICS RELATED TO TRANSFORMATIONS, DISTRIBUTION FUNCTIONS AND STOCHASTIC PROCESSES

NORMAN C. SEVERO
PAUL J. SCHILLO
ROBERT H. RODINE

STATE UNIVERSITY OF NEW YORK AT BUFFALO

NOVEMBER 1966

Contract AF 33(657)-9885
Project 7071

Distribution of this document is unlimited

AEROSPACE RESEARCH LABORATORIES
OFFICE OF AEROSPACE RESEARCH
UNITED STATES AIR FORCE
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

[redacted]

FOREWARD

This final technical report was prepared by the State University of New York at Buffalo on Contract AF 33(657)-9885 for the Aerospace Research Laboratories, United States Air Force, under the technical cognizance of Dr. H. Leon Harter of the Applied Mathematics Research Laboratory of ARL.

The object of the report is to investigate some topics related to measuring the effectiveness of transformations used in mathematical and applied statistics.

The report is divided into four parts with authorship as follows: Part I, R. H. Rodine; Part II, P. J. Schillo; Part III, R. H. Rodine, P. J. Schillo, and N. C. Severo; and Part IV, N. C. Severo.

The authors wish to thank Dr. Harter for his continued interest in this project and for his constructive criticism.

ABSTRACT

The purpose of Section I is to generalize Theorem 3 of Severo, Montzingo, and Schillo, "Characterization of the asymptotic distributions of a transformed normal random variable," Sankhyā, Series A, 27, by relaxing the assumption of normality and removing the requirement that the parameters μ_y and σ_y of the distribution of the random variable Y be the mean and variance, respectively. A result analogous to the above-mentioned theorem is obtained for the class of location-scale parameter distributions. Examples are given which show that the conditions given are sufficient, but not necessary, for the existence of an asymptotic distribution of a transformed random variable.

Section II illustrates ways of deciding whether or not a given univariate random variable X can be transformed into a given univariate random variable Y ; and it gives procedures for defining various transformations of X into Y in the event that one such transformation is known to exist.

Section III, which consists of three parts, gives an illustration of the notion of robustness of a test, a generalization of this notion, and a tentative definition of the robustness of a test in terms of a metric on the space of power functions of the test. A short investigation is also made in this section of the properties of the Kolmogorov metric on the space of location families of distribution functions, and applications are made to the normal, Cauchy, student's-t, gamma, and exponential distributions.

Section IV presents two theorems that provide simple interative solutions of special systems of differential-difference equations. The first system consists of linear differential equations whose coefficient matrices are triangular, have constant elements, and have diagonal elements equal to each other at most in pairs. The equations of the second system also have constant triangular coefficient matrices, such that whenever there are equal diagonal elements then sufficient conditions are imposed on the matrices themselves so that the solutions involve only sums of exponential terms.

The theorems are applied to the simple stochastic epidemic and to the general stochastic epidemic, respectively, in each of which the initial distribution of the number of uninfected susceptibles and the number of infectives are arbitrary but the total population size is assumed bounded. The results for the simple stochastic epidemic provide solutions not obtainable by previously known results. The results for the general stochastic epidemic are simpler and more direct than other known methods, which, when used to solve the problem having an arbitrary initial distribution, would involve additional steps that would sum proportionally-weighted conditional results.

Table of Contents

LIST OF ILLUSTRATIONS

I. ON THE ASYMPTOTIC DISTRIBUTION OF A TRANSFORMED RANDOM VARIABLE

In [2], there is stated and proved a theorem regarding the asymptotic distribution of a transformed normal random variable, and the following classes of Baire functions are defined:

For each natural number n, and each real number s, $\Lambda[n;s]$ is that class of Baire functions whose elements, λ, satisfy the conditions: (i) the n^{th} derivative, $\lambda^{(n)}$, of λ is continuous at s, (ii) $\lambda^{(n)}(s)$ is a nonzero real number, and (iii) if m is a natural number less than n, then $\lambda^{(m)}(s)$ is zero.

For each natural number n, and each real number s, $\Omega[n;s]$ is that class of Baire functions whose elements, ω, satisfy $\lim_{x\to 0} \omega(x)/x^n = s/n!$.

The theorem in [2] to which we refer may be stated in a slightly altered form as follows:

If Y is a normal random variable with mean μ_y and standard deviation σ_y ; if $Z = (Y-\mu_y)/\sigma_y$; if there is a natural number n such that $h \in \Lambda[n;\mu_y]$; if there are real constants R and Q, with $Q \neq 0$, such that $r \in \Omega[n;Rh^{(n)}(\mu_y)]$ and $\mu_0(\sigma_y) = h(\mu_y) + r(\sigma_y)$; if $\sigma_0 \in \Omega[n;Qh^{(n)}(\mu_y)]$; if $X = H(Y)$, and if $W = [X-\mu_0(\sigma_y)]/\sigma_0(\sigma_y)$; then the asymptotic distribution of W, as σ_y approaches zero, is that of the random variable $(Z^n-R)/Q$.

The purpose of the present work is to relax the assumption of normality on Y and, incidentally, to remove the requirement that the parameters μ_y and σ_y be the mean and standard deviation of Y .

For convenience, we introduce some terminology.

Let T be a subset of $R_1 \times R_1^+$, the real upper half plane. We call a family $\{F(x;\mu,\theta):(\mu,\theta) \in T\}$ of distribution functions a location-scale-parameter family with respect to T if, and only if, there is a distribution function, G, on the real line such that for each real x, we have $F(x;\mu,\theta) = G((x-\mu)/\theta)$, for every $(\mu,\theta) \in T$.

For ease of reference, we restate Lemma 2 of [2], in slightly altered form.

LEMMA. Let $h \in \Lambda[n;a]$, R be a real number, and Q be a nonzero real number. Then a necessary and sufficient condition that $\lim_{b\to 0} [h(a+bz) - h(a) - r(b)]/q(b) = (z^n-R)/Q$, for each real number z, is that $r \in \Omega[n;Rh^{(n)}(a)]$ and $q \in \Omega[n;Qh^{(n)}(a)]$.

We can now prove a generalized version of Theorem 3 of [2].

THEOREM. Let Y be a random variable with distribution function $F(y;\mu,\theta)$, where $\{F(y;\mu,\theta):(\mu,\theta) \in T\}$ is a location-scale family with respect to T. Let $h \in \Lambda[n;a]$, R be any real number, and Q be any nonzero real number. If $W = [h(Y) - h(\mu) - r(\theta)]/q(\theta)$ with $r \in \Omega[n;Rh^{(n)}(\mu)]$ and $q \in \Omega[n;Qh^{(n)}(\mu)]$, then the asymptotic distribution of W, as θ approaches zero, is equal to that of $(Z^n-R)/Q$, where $Z = (Y-\mu)/\theta$.

<u>Proof</u>. By applying Theorem a, page 166, of [1], the characteristic function, $\beta_W(t)$, may be written in the form

$$\beta_W(t) = \int_{-\infty}^{\infty} \exp[it\, w(z)]dG(z) = \int_{-\infty}^{\infty} \exp\{it[h(\mu+\theta z) - h(\mu) - r(\theta)]/q(\theta)\}dG(z) ,$$

where G, the distribution function of Z, does not depend on (μ,θ). By applying the Lebesgue dominated convergence theorem and the Lemma, we have, for each real t,

$$\lim_{\theta\to 0} \beta_W(t) = \int_{-\infty}^{\infty} \exp[it(z^n-R)/Q]dG(z) = \beta(t) \quad \text{, say.}$$

Since $\lim_{t\to 0} \beta(t) = \beta(0)$, β is continuous at zero and, hence, β is a characteristic function. Thus, the asymptotic distribution of W, as θ approaches zero, is equal to the distribution of $(Z^n-R)/Q$.

We note in passing that the proof does not depend upon the existence of moments of Y. For example, the theorem applies to the situation in which Y is a Cauchy random variable with location-scale parameter (μ,θ) belonging to $T = R_1 \times R_1^+$. Thus, for any transformation h of the type described in the theorem, the asymptotic distribution of the random variable $W = [h(Y) - h(\mu) - r(\theta)]/q(\theta)$ is equal to that of the random variable $(Z^n-R)/Q$, where the probability density function of Z is given by

$$f(z) = \frac{1}{\pi} \frac{1}{1+z^2} \quad , \quad -\infty < z < \infty \quad .$$

We also note that the theorem gives conditions under which the distribution of the random variable W approaches, as θ approaches zero, the distribution of the random variable $(Z^n-R)/Q$ which does not depend upon θ.

We give below some examples of situations in which the distribution of $(Z^n-R)/Q$ may depend upon θ, and for which it is true that the distribution functions of W and of $(Z^n-R)/Q$ approach the same function as θ approaches zero.

Example 1. In this example, we show that the distribution functions of W and of $(Z^n-R)/Q$ may approach the same distribution function as θ approaches zero, even though $F(y;\mu,\theta)$ is not a location-scale family with respect to T .

Let μ be a nonzero real number and let θ be a real number in the open unit interval $(0,1)$. Let $T' = \{(\mu,\theta)\}$ and let $Y(\mu,\theta)$ denote a discrete random variable with distribution function

$$F(y;\mu,\theta) = \begin{cases} 0 & , \ -\infty < y < \mu-\theta \\ \theta^2/2 & , \ \mu-\theta \le y < \mu \\ 1-\theta^2/2 & , \ \mu \le y < \mu+\theta \\ 1 & , \ \mu+\theta \le y < \infty \end{cases} .$$

If $Z(\mu,\theta)$ denotes the random variable $(Y(\mu,\theta)-\mu)/\theta$, then the distribution function of $Z(\mu,\theta)$ is given by

$$G(z;\theta) = \begin{cases} 0 & , \ -\infty < z < -1 \\ \theta^2/2 & , \ -1 \le z < 0 \\ 1-\theta^2/2 & , \ 0 \le z < 1 \\ 1 & , \ 1 \le z < \infty \end{cases} ,$$

which depends on (μ,θ) , so that $\{F(y;\mu,\theta) : (\mu,\theta) \in T'\}$ is not a location-scale family with respect to T' .

If R is any real number and Q is any positive real number, then we denote by $U(\theta)$ the random variable $(Z^2(\mu,\theta)-R)/Q$ in order to stress the fact that the distribution function of $Z(\mu,\theta)$ depends on θ . It is easy to show that the distribution function of $U(\theta)$ is given by

$$H(u;\theta) = \begin{cases} 0 & , \quad -\infty < u < -R/Q \\ 1-\theta^2 & , \quad -R/Q \le u < (1-R)/Q \\ 1 & , \quad (1-R)/Q \le u < \infty \end{cases} .$$

If we take $h(y) = (y-\mu)^2$, $r(\theta) = R\theta^2(1+\theta)$ and $q(\theta) = Q\theta^2(1+2\theta)$, then it is readily verified that $h \in \Lambda[2;\mu]$, $r \in \Omega[2;2R]$ and $q \in \Omega[2;2Q]$, and that the random variable $W(\theta)$ defined by $W(\theta) = [h(Y(\mu,\theta)) - h(\mu) - r(\theta)]/q(\theta)$ has distribution function

$$J(w;\theta) = \begin{cases} 0 & , \quad -\infty < w < -R(1+\theta)/Q(1+2\theta) \\ 1-\theta^2 & , \quad -R(1+\theta)/Q(1+2\theta) \le w < [1-R(1+\theta)]/Q(1+2\theta) \\ 1 & , \quad [1-R(1+\theta)]/Q(1+2\theta) \le w < \infty \end{cases} .$$

If we consider, for a fixed real number x , the limit, as θ approaches zero, of $H(x;\theta)$, then we see that

$$\lim_{\theta \to 0} H(x;\theta) = \begin{cases} 0 & , \quad -\infty < x < -R/Q \\ 1 & , \quad -R/Q \le x < \infty \end{cases} .$$

And if we write $J(w;\theta)$ in the equivalent, but more tractable, form

$$J(w;\theta) = \begin{cases} 0 & , \quad -\infty < w < -R/Q\left[1 + \frac{\theta}{1+\theta}\right]^{-1} \\ 1-\theta^2 & , \quad -R/Q\left[1 + \frac{\theta}{1+\theta}\right]^{-1} \le w < [Q(1+2\theta)]^{-1} - R/Q\left[1 + \frac{\theta}{1+\theta}\right]^{-1} \\ 1 & , \quad [Q(1+2\theta)]^{-1} - R/Q\left[1 + \frac{\theta}{1+\theta}\right]^{-1} \le w < \infty \end{cases} ,$$

then it is easy to see that, for fixed real x , we have

$$\lim_{\theta \to 0} J(x;\theta) = \begin{cases} 0 & , \quad -\infty < x < -R/Q \\ 1 & , \quad -R/Q \le x < \infty \end{cases} .$$

Thus, as θ approaches zero, the distribution functions of the random variables $W(\theta)$ and $U(\theta)$ both approach the same distribution function.

Example 2. In this example, we show that even though the limit, as θ approaches zero, of the distribution function of the random variable $U(\theta) = (Z^n(\mu,\theta)-R)/Q$ is not a distribution function, the limit, as θ approaches zero, of the difference between the distribution function of $U(\theta)$ and that of $W(\theta)$ can approach zero.

Let T' be the same as in example 1, and let $Y(\mu,\theta)$ denote a discrete random variable with distribution function

$$F(y;\mu,\theta) = \begin{cases} 0 & , -\infty < y < \mu-\theta^{-1} \\ (1/2)(1-\theta^2) & , \mu-\theta^{-1} \le y < \mu \\ (1/2)(1+\theta^2) & , \mu \le y < \mu+\theta^{-1} \\ 1 & , \mu+\theta^{-1} \le y < \infty \end{cases} .$$

The distribution function of the random variable $(Y(\mu,\theta)-\mu)/\theta$, which we again denote by $Z(\mu,\theta)$, is given by

$$G(z;\theta) = \begin{cases} 0 & , -\infty < z < -\theta^{-2} \\ (1/2)(1-\theta^2) & , -\theta^{-2} \le z < 0 \\ (1/2)(1+\theta^2) & , 0 \le z < \theta^{-2} \\ 1 & , \theta^2 \le z < \infty \end{cases} ,$$

and that of the random variable $U(\theta) = (Z^2(\mu,\theta)-R)/Q$, for positive real R and positive real Q, is given by

$$H(u;\theta) = \begin{cases} 0 & , -\infty < u < -R/Q \\ \theta^2 & , -R/Q \le u < (\theta^{-4}-R)/Q \\ 1 & , (\theta^{-4}-R)/Q \le u < \infty \end{cases} .$$

By taking h , r and q to be the same functions as they were in example 1 and defining $W(\theta)$ as it was defined there, we find that $W(\theta)$ has distribution function

$$J(w;\theta) = \begin{cases} 0 & , -\infty < w < -R(1+\theta)/Q(1+2\theta) \\ \theta^2 & , -R(1+\theta)/Q(1+2\theta) \le w < [\theta^{-4}-R(1+\theta)]/Q(1+2\theta) \\ 1 & , [\theta^{-4}-R(1+\theta)]/Q(1+2\theta) \le w < \infty \end{cases} .$$

In order to simplify the computations and make the discussion easier to follow, we now specialize to the case $R = 1$ and $Q = 1$. We note that no essential generality is lost by so doing. Under this restriction we have

$$H(u;\theta) = \begin{cases} 0 & , -\infty < u < -1 \\ \theta^2 & , -1 \le u < \theta^{-4}-1 \\ 1 & , \theta^{-4}-1 \le u < \infty \end{cases} ,$$

and

$$J(w;\theta) = \begin{cases} 0 & , -\infty < w < -[1+\theta/(1+\theta)]^{-1} \\ \theta^2 & , -[1+\theta/(1+\theta)]^{-1} \le w < [\theta^{-4}-(1+\theta)]/(1+2\theta) \\ 1 & , [\theta^{-4}-(1+\theta)]/(1+2\theta) \le w < \infty \end{cases} .$$

It is clear that, for any fixed real number x , the absolute difference $|J(x;\theta) - H(x;\theta)|$ can be made arbitrarily small by choosing θ sufficiently close to zero; i.e.,

$$\lim_{\theta \to 0} |J(x;\theta) - H(x;\theta)| = 0 \quad , \quad -\infty < x < \infty .$$

It is also clear that if x is any fixed real number, then $H(x;\theta)$ can be made arbitrarily small by choosing θ sufficiently close to zero (for $\theta < (x+1)^{-\frac{1}{4}}$, $H(x;\theta) \leq (x+1)^{-\frac{1}{8}}$); i.e.,

$$\lim_{\theta \to 0} H(x;\theta) = 0 \quad , \quad -\infty < x < \infty \quad ,$$

which is not a distribution function.

REFERENCES

[1] Loeve, M. (1963), Probability Theory, third edition, van Nostrand, Princeton.

[2] Severo, N. C., Montzingo, L. J., and Schillo, P. J. (1965), Characterization of the asymptotic distributions of a transformed normal random variable, Sankhyā, Series A, 27, 417-422.

II. SUPPLEMENT TO CHAPTER II OF ARL 65-75

1. Preliminary Remarks

Let B denote the set of all Baire functions which map the real line R into R ; and let D denote that subset of B which contains the function F if, and only if, F is a distribution function. If F is a function in D , if P is the probability measure which induces F , so that, at each point x of R , the value of F is

$$F(x) = P\left\{ z : z \leq x \right\} ,$$

if h is a function in B , and if G is that function in D whose value at each point y of R is

$$G(y) = P\left\{ x : h(x) \leq y \right\} ,$$

then the symbol (F,h,G) stands for the assertion that the Baire functions F , h and G are related to one another in the manner hypothesized. For each ordered pair (F,G) of functions in D , the symbol (F,*,G) denotes the subset $\left\{ h : (F,h,G) \right\}$ of B ; and (F,G) is said to be a <u>compatible</u> pair or an <u>incompatible</u> pair according as it is not or is true that the set (F,*,G) is the empty set $\emptyset$. Chapter II of the interim technical report ARL 65-75 , entitled "Some Properties Of Distribution Functions And Transformations That Induce One Another", provides necessary and sufficient conditions under which an ordered pair (F,G) of distribution functions in D is a compatible pair. In this supplement to chapter II of ARL 65-75 , we shall carefully examine some special ordered pairs of distribution functions in D ; we shall determine whether or not they are compatible pairs; if one of them, (F,G) , is a compatible pair, then we shall determine at least one function h in the non-empty set $(F,*,G)$; and, in the final section of this supplement, we shall consider ways of finding various functions in $(F,*,G)$ when (F,G) is a compatible pair. We would like our illustrations to require -- and, thereby, to justify -- some of the elaborate details in the theory concerning compatible pairs of functions in D which is developed in chapter II of ARL 65-75.

In order to make this supplement to chapter II of ARL 65-73 notationally independent of that technical report, we shall define here those of its symbols and terms that we use. Of course, many statements in this supplement will not be independent of that reference, because we shall not prove them here.

2. The Notion of Clearance

Let W denote the set of all infinite column matrices whose entries are real numbers. Thus, W may also be regarded as the set of all infinite sequences of real numbers. If w denotes a matrix in W (or a sequence in W), then, for each positive integer n , the symbol w_n denotes the entry in the n'th row of w (or the n'th term of w).

Three operators which may be applied to symbols denoting elements of W are cum: , lim: and sum: . If w denotes an element of W , then cum:w denotes the sequence of partial sums of the terms of the infinite sequence w ; that is, cum:w denotes that element v of W whose first term v_1 is w_1 , and whose n'th term v_n , for each integer $n > 1$, is $w_n + v_{n-1}$. If w denotes a convergent sequence in W , then lim:w denotes the limit of w . Finally, if w denotes an element of W which is such that cum:w is a convergent sequence, then sum:w denotes the real number lim:cum:w .

A matrix v in W is said to <u>dominate</u> a matrix w in W if, and only if, for each positive integer n , $v_n \geq w_n$; and, the fact that v dominates w may be asserted symbolically by either $v \gg w$ or $w \ll v$. If $v \gg w$ and $v \neq w$, then we may write $v \gg w$ and $w \ll v$.

Let $\underline{T}$ denote the set of all matrices with infinitely many rows and columns, in each column of which there is one, and only one, non-zero entry, the real number 1 . If A denotes a matrix in $\underline{T}$, then, for each ordered pair m,n of positive integers, the symbol $A_{m,n}$ denotes the entry in the m'th row and n'th column of A . Let $\underline{I}$ denote the identity matrix in $\underline{T}$; that is, let $\underline{I}$ be that matrix in $\underline{T}$ which is such that, for each positive integer n , $\underline{I}_{n,n} = 1$.

For each matrix w in W , let $\underline{T}(w)$ denote that subset of $\underline{T}$ which contains the matrix A in $\underline{T}$ if, and only if, the matrix product Aw is a defined column matrix; in other words, let $\underline{T}(w)$ denote that subset of $\underline{T}$ which contains the matrix A in $\underline{T}$ if, and only if, for each positive integer m , the sequence v in W , whose n'th term v_n is $A_{m,n}w_n$ for each positive integer n , is such that sum:v is defined. It is evident that, for each matrix w in W , $\underline{T}(w) \neq \emptyset$, because $\underline{I}w = w$.

A matrix v in W is said to <u>clear</u> a matrix w in W if, and only if, there is at least one matrix A in $\underline{T}(w)$ for which it is true that v dominates the matrix product Aw ; and, the fact that v clears w may be asserted symbolically by either $v \ggg w$ or $w \lll v$.

Our special interest in the notion of clearance -- that is, the idea which underlies the assertion that one infinite sequence of real numbers clears another such sequence -- will be restricted to such a modest range of its application (e.g., to such sequences in W as w , where w is a monotone non-increasing sequence of non-negative real numbers, and where sum:w is a definite non-negative real number that does not exceed 1) that it might seem that some simpler notion would serve our purpose for introducing it. It happens, however, that we have found no replacement for it that is as easily described as it is, and that is suitable to our needs.

3. The Theorem in Chapter II of ARL 65-75

Some notation from chapter II of ARL 65-75 is the following: For each real number x , $\bar{\omega}(x)$ denotes that subset of R which contains the point z of R if, and only if, $z \leq x$. For each function F in D , the symbol $R(F;\infty)$ denotes the set of all points of discontinuity of F ; the symbol $R'(F;\infty)$ denotes the difference set $R - R(F;\infty)$; the symbol $R_c(F;\infty)$ denotes that subset of $R'(F;\infty)$ which contains the point x of $R'(F;\infty)$ if, and only if, for each point z of R which is less than x , $F(z) < F(x)$; the symbol $R_c'(F;\infty)$ denotes the difference set $R'(F;\infty) - R_c(F;\infty)$; the symbol $R_w(F;\infty)$ denotes that subset of $R_c(F;\infty)$ which contains the point x of $R_c(F;\infty)$ if, and only if, for each real number $z < x$, and with P denoting the probability measure that induces F ,

$$P\left\{ r : r \in \bar{\omega}(z) \cap R_c(F;\infty) \right\} < P\left\{ r : r \in \bar{\omega}(x) \cap R_c(F;\infty) \right\} ;$$

and the symbol $R_w'(F;\infty)$ denotes the difference set $R'(F;\infty) - R_w(F;\infty)$.

Finally, for the function F in D and the real number x, the symbols $R(F;x)$, $R'(F;x)$, $R_c(F;x)$, $R_c'(F;x)$, $R_w(F;x)$ and $R_w'(F;x)$ denote the set intersections $\bar{\omega}(x) \cap R(F;\infty)$, $\bar{\omega}(x) \cap R'(F;\infty)$, $\bar{\omega}(x) \cap R_c(F;\infty)$, $\bar{\omega}(x) \cap R_c'(F;\infty)$, $\bar{\omega}(x) \cap R_w(F;\infty)$ and $\bar{\omega}(x) \cap R_w'(F;\infty)$, respectively.

Let ρ be that function which maps D into W in such a way that its value at each function F in D is that sequence $\rho(F) = w$ in W which satisfies the following six conditions:

(1) if $R(F;\infty) = \emptyset$, then, for each positive integer n, $w_n = n$;

(2) if $R(F;\infty)$ contains exactly m points, the greatest of which is r, then, for each integer $n > m$, $w_n = r + n - m$;

(3) if a is a point of $R(F;\infty)$, then there is one, and only one, positive integer n such that $w_n = a$;

(4) if, for some positive integer n, w_n is not a point of $R(F;\infty)$, then $R(F;\infty)$ is not an infinite set;

(5) if a and b are points of $R(F;\infty)$ such that the saltus of F at a is greater than the saltus of F at b, then there exist positive integers m and n such that $m < n$, $w_m = a$ and $w_n = b$; and

(6) if a and b are points of $R(F;\infty)$ such that $a < b$, and such that the saltus of F at a is the same as the saltus of F at b, then there exist positive integers m and n such that $m < n$, $w_m = a$ and $w_n = b$.

Let W' be that subset of W which contains the sequence w of W if, and only if, cum:w is a convergent, monotone non-decreasing sequence, w is a monotone non-increasing sequence, and sum:w does not exceed 1.

If F is any function in D, let w be the sequence $\rho(F)$ in W; for each positive integer m, let $v^{(m)}$ be that sequence in W whose n'th term, for each positive integer n, is $v_n^{(m)} = F(w_m) - F(w_m - 1/n)$; and let the symbol sal:F denote that sequence u in W' whose m'th term, for each positive integer m, is $u_m = \lim :v^{(m)}$.
Thus, sal: is an operator which is applicable to each function F in D; and, loosely speaking, sal:F is the monotone non-increasing sequence whose non-zero terms are the saltuses of F.

By making use of the notation described in this supplement to chapter II of ARL 65-75, the theorem of that chapter can be stated concisely as follows:

For F and G in D, $(F,*,G) \neq \emptyset$ if, and only if, sal:F $\lll$ sal:G.
Thus, this theorem gives necessary and sufficient conditions for an ordered pair (F,G) of functions in D to be a compatible pair. However, since these conditions appeal to the somewhat formidable notion of clearance for their meanings, it behooves us to show that this theorem is necessarily preferable to the tautological assertion that (F,G) is a compatible pair if, and only if, (F,G) is a compatible pair. In the next section, we intend to provide the theorem with a modest justification.

4. Clearance in W'

In this section, we shall state and prove three theorems concerning the clearance relation $\ggg$ between sequences in W'. These theorems will be considered in the next section where we shall be dealing with some special compatible and incompatible pairs of discrete distribution functions in D. However, we would not want the theorems of this section to be mistaken for fundamental theorems in a carefully constructed theory of clearance. Such a study would properly begin on foundations provided by a good understanding of the set W , and of the <u>dominance relation</u> $\gg$ between its elements, and by a separate study of the set $\underline{T}$. In unifying the results of these separate investigations in order to develop a theory of clearance, one of the fundamental theorems would assert that clearance, like dominance, is a transitive relation. The following theorems are too limited in scope to be regarded as fundamental theorems.

THEOREM 1. If the sequence v in W' dominates the sequence w in W' , then v clears w .

PROOF: Since $v \gg w$, $\underline{I} \in \underline{T}(w)$ and $\underline{I}w = w$, $v \gg \underline{I}w$; so $v \ggg w$.

THEOREM 2. The sequence v in W' clears the sequence w in W' only if cum:v dominates cum:w .

PROOF: We shall prove this theorem in the following way: we shall let v and w be any sequences in W' which are such that the sequence y = cum:v in W does not dominate the sequence x = cum:w in W ; we shall let A be a matrix in $\underline{T}(w)$; and we shall let u be the matrix product Aw in W' ; then we shall show that A cannot meet the necessary requirements for v to dominate u , so that v cannot clear w .

Since y does not dominate x , there exists one, and only one, positive integer k such that $y_k < x_k$, and such that, if n is a positive integer less than k , then $y_n \geq x_n$. Since v and w are monotone non-increasing sequences of non-negative real numbers, it follows that

(*) for any positive integers m and n such that $m \geq k \geq n$, $v_m < w_n$.

For each positive integer n , let $r(A;n)$ denote that positive integer m which is such that $A_{m,n} = 1$; and, for each positive integer m , let $A^{(m)}$ denote that sequence in W whose n'th term $A_n^{(m)}$ is $A_{m,n}$ for each positive integer n . Now, if, for some positive integer $n \leq k$, the integer $m = r(A;n)$ is not less than k , then, by (*) ,

$$u_m = \sum_{h=1}^{\infty} A_h^{(m)} w_h \geq A_n^{(m)} w_n = w_n > v_m \quad ,$$

so that v cannot dominate $u = Aw$; therefore, if v is to clear w , A must be such a matrix in $\underline{T}$ that, for each positive integer $n \leq k$, $r(A;n) < k$. If A meets this requirement, and if z denotes the sequence cum:u in W , then

$$z_k = \sum_{n=1}^{k} u_n = \sum_{n=1}^{k} \sum_{h=1}^{\infty} A_h^{(n)} w_h \geq \sum_{n=1}^{k} \sum_{h=1}^{k} A_h^{(n)} w_h = \sum_{h=1}^{k} w_h \sum_{n=1}^{k} A_h^{(n)}$$
$$= \sum_{h=1}^{k} w_h A_h^{(r(A;h))} = \sum_{h=1}^{k} w_h = x_k \quad ;$$

and, since $x_k > y_k$, $z_k > y_k$. Consequently, there is a positive integer $n \leq k$ such that $u_n > v_n$, so that v does not dominate u . Thus, for no matrix A in $\underline{T}(w)$, does v dominate Aw ; hence, v does not clear w .

THEOREM 3. The sequence v in W' clears the sequence w in W' only if $\text{sum}:v \geq \text{sum}:w$.

PROOF: Let $y = cum:v$ and $x = cum:w$. If $sum:v < sum:w$, then, because v and w are monotone non-increasing sequences of non-negative real numbers, there is a positive integer k such that $x_k > sum:v$; hence, $x_k > y_k$, so that cum:v does not dominate cum:w and, by theorem 2, v does not clear w .

5. Examples of Compatible and Incompatible Pairs

Let I_0 denote the negative half of the real line R ; let J_0 denote the non-negative half of R ; and, for each positive integer n , let I_n , I'_n , J_n and J'_n denote the respective intervals $[\, n - 1 , n - 1/2 \,)$, $[\, n - 1/2 , n \,)$, $[\, -2n , -2n + 99/100 \,)$ and $[\, -2n + 99/100 , -2n + 2 \,)$, which are all open on the right.

Let F be that function in D whose value at each $x \in R$ is that real number $F(x)$ which is defined as follows: if $x \in I_0$, then $F(x) = 101/2(101-x)$; if n is a positive integer and $x \in I_n$, then $F(x) = (101x+2n^2+200n+10100)/2(n+100)(n+101)$; and, if n is a positive integer and $x \in I'_n$, then $F(x) = (4n^2+602n+20099)/4(n+100)(n+101)$.

Let H be that function in D whose value at each $x \in R$ is that real number $H(x)$ which is defined as follows: if $x \in I_0$, then $H(x) = 50/(100-x)$; if n is a positive integer and $x \in I_n$, then $H(x) = (50x+n^2+99n+4950)/(n+99)(n+100)$; and, if n is a positive integer and $x \in I'_n$, then $H(x) = (n^2+149n+4925)/(n+99)(n+100)$.

Let G be that function in D whose value at each $x \in R$ is that real number $G(x)$ which is defined as follows: if $x \in J_0$, then $G(x) = (x+1)/(x+2)$; if n is a positive integer and $x \in J_n$, then $G(x) = (x+4n)/4n(n+1)$; and, if n is a positive integer and $x \in J'_n$, then $G(x) = (200n+99)/400n(n+1)$.

We intend to find out whether or not (F,H) and (F,G) are compatible pairs. Let $\text{sal}:F = w$, $\text{sal}:H = u$, $\text{sal}:G = v$, $\text{cum}:w = \bar{w}$, $\text{cum}:u = \bar{u}$ and $\text{cum}:v = \bar{v}$; and let a , b , $\bar{a}$ and $\bar{b}$ be those sequences in W whose n'th terms, for each positive integer n , are $a_n = u_n - w_n$, $b_n = v_n - w_n$, $\bar{a}_n = \bar{u}_n - \bar{w}_n$ and $\bar{b}_n = \bar{v}_n - \bar{w}_n$, respectively.

We shall consider the pair (F,H) first. Each of the sets $R(F;\infty)$ and $R(H;\infty)$ is the set of all positive integers; in fact, $\rho(F)$ and $\rho(H)$ are the sequence of positive integers. Furthermore, for each positive integer n , $w_n = 101/4(n+100)(n+101)$, $u_n = 25/(n+99)(n+100)$, $a_n = (101-n)/4(n+99)(n+100)(n+101)$, $\bar{w}_n = n/4(n+101)$, $\bar{u}_n = n/4(n+100)$ and $\bar{a}_n = n/4(n+100)(n+101)$. Thus, $\text{sum}:w = \text{lim}:\bar{w} = 1/4$ and $\text{sum}:u = 1/4$. In theorems 2 and 3 of section 4 , there were given necessary conditions for u to clear w ; since $\bar{u}$ dominates $\bar{w}$ and $\text{sum}:w$ does not exceed $\text{sum}:u$, these necessary conditions are met; and, therefore, it cannot be concluded that (F,H) is an incompatible pair by applying those theorems. In theorem 1 of section 4 , there was given a sufficient condition for u to clear w ; since $a_{102} < 0$, so that $u_{102} < w_{102}$ and u does not dominate w , this sufficient condition is not met; and, therefore, it cannot be concluded that (F,H) is a compatible pair by applying that theorem. In this example, the simple tests of section 4 are of no use to us. We have to find some other way of determining whether or not the set $(F,*,H)$ is empty.

We designed this problem to illustrate some of the difficulties which can attend the use of the clearance criterion in testing the compatibility of ordered pairs of distribution functions in D . However, since we wanted a decision to come out of a modest amount of study, we made it possible to show that (F,H) is an incompatible pair (i.e., that u does not clear w) in a way that is only a little more complicated than it would have been if the tests derived from the theorems of section 4 could have been employed successfully. Of course, in this way, we do no more than hint that there might be ordered pairs of functions in D to which neither the adjective "compatible" nor the adjective "incompatible" can be applied. Until an upper bound on the possible number of simple decisions necessary to conclude that a given ordered pair of functions in D is or is not a compatible pair is definitely established, the statement that the pair has to be either compatible or incompatible will appear to be quite flimsy from a critical viewpoint. We must admit that there is much that is vulnerable to severe criticism in our simple dichotomy of the Cartesian product set $D \times D$ of all ordered pairs of functions in D into compatible and incompatible pairs; but no example here is subject to that kind of criticism.

In showing that (F,H) is an incompatible pair, our procedure will be as follows: we shall try to find a matrix A in $\underline{T}(w)$ which is such that, if z is the matrix product Aw in W' , then u dominates z ; therefore, when we have shown that no such matrix A can meet all the requirements which we shall find that we must impose on it, we shall have shown that u cannot clear w , and that (F,H) is an incompatible pair.

For each positive integer n , let $r(A;n)$ denote that positive integer m for which $A_{m,n} = 1$. For each positive integer n , $w_n > u_{n+1}$, because $w_n - u_{n+1} = 1/4(n+100)(n+101)$; therefore, if $r(A;n)$ were to be an integer m greater than n , then the m'th term of the sequence

$z = Aw$ would be

$$z_m = \sum_{t=1}^{\infty} A_{m,t} w_t \geq A_{m,n} w_n = w_n > u_{n+1} \geq u_m \quad ,$$

in which event u would not dominate z ; consequently, A must meet the requirement that, for each positive integer n , $r(A;n) \leq n$.

Since $r(A;1) \leq 1$, $r(A;1) = 1$. Suppose that it has been shown that, for each positive integer n which is less than or equal to a particular positive integer k , $r(A;n)$ must be n if it is to be at all possible for u to dominate $z = Aw$. We wish to see whether or not, under these conditions, $r(A;k+1)$, which cannot exceed $k+1$, can be a positive integer m which is less than $k+1$. For such an integer m , the m'th term of z would be

$$z_m = \sum_{t=1}^{\infty} A_{m,t} w_t \geq A_{m,m} w_m + A_{m,k+1} = w_m + w_{k+1} = u_m + (w_{k+1} - a_m) \quad .$$

If w_{k+1} is greater than a_m , then z_m would be greater than u_m , and u would not dominate z . Hence $r(A;k+1)$ cannot be an integer m less than $k+1$ if w_{k+1} exceeds a_m . Since, for each positive integer $m < 201$, $a_m - a_{m+1} = (201-m)/2(m+99)(m+100)(m+101)(m+102)$ is positive, and since, for each integer $m > 101$, $a_m = (101-m)/4(m+99)(m+100)(m+101)$ is negative, the greatest term in the sequence a of W is $a_1 = 1/41208$. Therefore, since w_{k+1} must not exceed $a_m \leq a_1$ if it is to remain possible for m to be less than $k+1$, it follows that m cannot be less than $k+1$ if w_{k+1} exceeds a_1 . We find that the least positive integer k for which it is true that w_{k+1} does not exceed a_1 is the least positive integer k

such that $101/4(k+101)(k+102) \leq 1/41208$; thus, we conclude that k must exceed $(\sqrt{4162009} - 203)/2$, which means that $k \geq 919$. For this reason, we are forced to conclude that, if k were to be less than 919 , then $r(A;k+1)$ would have to be $k+1$; consequently, for each positive integer $n < 920$, A must meet the requirement that $r(A;n) = n$. However, if A meets this requirement, then, since $a_{102} < 0$, the 102'nd term of $z = Aw$ is

$$z_{102} = \sum_{t=1}^{\infty} A_{102,t} w_t \geq A_{102,102} w_{102} = w_{102} = u_{102} - a_{102} > u_{102} ,$$

so that u does not dominate z . Thus, no matrix A in $\underline{T}(w)$ is such that u dominates Aw ; consequently, u does not clear w ; and (F,H) is an incompatible pair.

Now, let us consider the ordered pair (F,G) . It will not be hard for us to show that (F,G) is a compatible pair, so that the subset $(F,*,G)$ of B is not $\emptyset$. Our work on this compatible pair will be primarily that of defining a particular Baire function h in $(F,*,G)$. But, first, we shall take the trouble to show that $sal{:}G = v$ clears $sal{:}F = w$.

We find that $R(G;\infty)$ is the set of all non-positive even integers, that $\rho(G)$ is the strictly monotone decreasing sequence of non-positive even integers, and that the n'th terms of the sequences v , $\bar{v}$, b and $\bar{b}$, for each positive integer n , are $v_n = 101/400n(n+1)$, $\bar{v}_n = 101n/400(n+1)$, $b_n = v_n - w_n = 101(10100+101n-99n^2)/400n(n+1)(n+100)(n+101)$ and $\bar{b}_n = \bar{v}_n - \bar{w}_n = n(n+10101)/400(n+1)(n+101)$, respectively, so that $sum{:}v = lim{:}\bar{v} = 101/400$. Evidently, the tests derived from the three theorems of section 4 fail to disclose whether or not (F,G) is a compatible pair.

We shall show that v clears w by defining a matrix A in $\underline{T}(w)$ which is such that v dominates Aw. This satisfactory matrix A will then be of use to us in our work defining a function h in $(F,*,G)$. For each ordered pair of positive integers (m,n), let $A_{m,n}$ be either 1 or 0 according as it is or is not true that $101(m-1) < n \leq 101m$. If z is the matrix product Aw in W', then, for each positive integer m, the m'th term of z is

$$z_m = \sum_{n=1}^{\infty} A_{m,n} w_n = \sum_{n=101m-100}^{101m} w_n = \frac{1}{4m(m+1)} < \frac{101}{400m(m+1)} = v_m ;$$

consequently, v dominates z, v clears w, and (F,G) is a compatible pair.

The matrix A establishes the following one-to-many correspondence between the points of $R(G;\infty)$ and $R(F;\infty)$; to each positive integer m, and, hence, to the m'th term $y_m = 2 - 2m$ of $\rho(G)$, there correspond 101 points of $R(F;\infty)$, the n'th of which is $x_{m,n} = 101(m-1) + n$ of $R(F;\infty)$ for each positive integer $n \leq 101$; and, to each positive integer n, and, hence, to the n'th term n of $\rho(F)$, there corresponds one, and only one, point of $R(G;\infty)$, the even integer $2y$ which is such that $(1-n)/101 \leq y \leq (101-n)/101$. We are now prepared to begin defining a function h in $(F,*,G)$ by following the procedure used in chapter II of ARL 65-75.

For any subsets M and N of the real line R, let the symbol $B[M,N]$ denote that subset of B which contains the Baire function q in B if, and only if, the value of q at each point x of M is a point

$q(x)$ of N. Let γ denote that special function in B whose value at each point x of R is that integer $\gamma(x)$ which is such that $x - 1 < \gamma(x) \leq x$. Since the function h in $(F,*,G)$ which we want to define is a function in $B[R(F;\infty),R(G;\infty)]$, we make use of the many-to-one correspondence between the points of $R(F;\infty)$ and $R(G;\infty)$ which was defined in the last paragraph; and we define the value of h at each point x of $R(F;\infty)$ to be

$$(1) \qquad h(x) = -2\,\gamma\left(\frac{4x-1}{404}\right) .$$

Of course, if we were to set up a different many-to-one correspondence between the points of $R(F;\infty)$ and the points of a subset of $R(G;\infty)$ (i.e., if we were to find a matrix $C \neq A$ in $\underline{T}(w)$ such that $v \underline{\gg} Cw)$, then we would have to define h differently over $R(F;\infty)$.

The subsets $R'_c(F;\infty)$ and $R'_w(F;\infty)$ of R are the same set; this set contains the point x of R if, and only if, there is a positive integer n such that I'_n contains x, and such that n is not $x + 1/2$. For each point x of $R'_c(F;\infty) = R'_w(F;\infty)$, there is a real point $y > x$ such that $F(y) = F(x)$. We are free to define h over $R'_w(F;\infty)$ in any way that we find convenient; therefore, we define the value of h at each point x of $R'_w(F;\infty)$ to be

$$(2) \qquad h(x) = -2\,\gamma\left(\frac{4x-1}{404}\right) .$$

It remains for us to define h over $R_w(F;\infty) = R'(F;\infty) - R'_w(F;\infty)$. Let p be that sequence in W' whose n'th term, for each positive integer

n, is $p_n = v_n - s_n = 1/400n(n+1)$; then the special number $t = \text{sum}:p$ is $1/400$. Let P denote the probability measure which induces the distribution function F ; and let T and S be Borel sets in R which are such that $T \cap S = \emptyset$, $T \cup S = R_w(F;\infty)$ and $P[T] = t$. Since $R = R(F;\infty) \cup R'(F;\infty)$ and $R'(F;\infty) = R_w'(F;\infty) \cup R_w(F;\infty) = R_w'(F;\infty) \cup T \cup S$, since $T \cap S = \emptyset$ and $R(F;\infty) \cap R_w'(F;\infty) = R(F;\infty) \cap T = R(F;\infty) \cap S = R_w'(F;\infty) \cap T = R_w'(F;\infty) \cap S = \emptyset$, and since $P[R] = 1$, $P[R(F;\infty)] = \text{sum}:w = 1/4$, $P[R_w'(F;\infty)] = 0$, and $P[T] = t = 1/400$, it follows that $R = R(F;\infty) \cup R_w'(F;\infty) \cup T \cup S$ and $P[R] = P[R(F;\infty)] + P[R_w'(F;\infty)] + P[T] + P[S]$, so that the special number $s = P[S]$ is $299/400$. There are non-denumerably infinitely many ways of choosing the disjoint Borel sets T and S so that $T \cup S = R_w(F;\infty)$, $P[T] = t$ and $P[S] = s$; and no two such choices would yield the same definition of h over $R_w(F;\infty)$. We shall define S (and, consequently, T) in a way that we feel is most convenient.

Let S_0 and S_0' denote the set of all non-positive real numbers; and, for each positive integer m, let S_m denote the open sub-interval $(m - 1 , m - 101/200)$ in I_m, and let

$$S_m' = \bigcup_{i=0}^{m} S_i .$$

Since, for any integers m and n such that $m > n \geq 0$, $S_m \cap S_n = \emptyset$, $S_m \subset R_w(F;m)$ and $S_n \subset R_w(F;n) \subset R_w(F;m)$, it follows that, for these integers, $P[S_m \cup S_n] = P[S_m] + P[S_n]$ and $S_n' \subset S_m' \subset R(F;m)$. Therefore, since $q_0 = P[S_0] = F(0) = 1/2$ and, for each positive integer m, $q_m = P[S_m] = F(m - 101/200) - F(m - 1) = 9999/400(m+100)(m+101)$, it follows that $q_0' = P[S_0'] = P[S_0] = 1/2$ and, for each positive integer m,

$$q'_m = P[S'_m] = P\left[\bigcup_{i=0}^{m} S_i\right] = \sum_{i=0}^{m} P[S_i] = \frac{1}{2} + \frac{99m}{400(m+101)} = \frac{299}{400} - \frac{9999}{400(m+101)} .$$

The facts that, if q' is that sequence in W' whose m'th term is q'_m for each positive integer m, then $\lim{:}q' = 299/400$, and that $S'_m \subset R_w(F;\infty)$ for each non-negative integer m, combine to enable us to define the set S to be the following union of Borel sets in $R_w(F;\infty)$:

$$\bigcup_{i=0}^{\infty} S_i .$$

For each positive integer m, let T''_m denote the closed interval $[m - 101/200 , m - 1/2]$ in R ; let

$$T_m = \bigcup_{i=1}^{101} T''_{101(m-1)+i} \text{ ; and let } T'_m = \bigcup_{i=1}^{m} T_i .$$

Since, for any integers m and n such that $m > n \geq 1$, $T''_m \cap T''_n = \emptyset$, $T''_m \subset R_w(F;m)$ and $T''_n \subset R_w(F;n) \subset R_w(F;m)$, it follows that, for these integers, $P[T''_m \cup T''_n] = P[T''_m] + P[T''_n]$, $T_m \subset R_w(F;101m)$ and $T_n \subset R_w(F;101n)$. Also, for any integers m and n such that $m > n \geq 1$, $T_m \cap T_n = \emptyset$, so that $P[T_m \cup T_n] = P[T_m] + P[T_n]$. Therefore, since, for each positive integer m, $p''_m = P[T''_m] = F(m - 1/2) - F(m - 101/200) = 101/400(m+100)(m+101)$, it follows that

$$p_m = P[T_m] = P\left[\bigcup_{i=1}^{101} T''_{101(m-1)+i}\right] = \sum_{i=1}^{101} P[T''_{101(m-1)+i}] = \frac{1}{400m(m+1)}$$

for each positive integer m. Furthermore, for each positive integer m,

$$p'_m = P[T'_m] = P\left[\bigcup_{i=1}^{m} T_i\right] = \sum_{i=1}^{m} P[T_i] = \frac{1}{400} - \frac{1}{400(m+1)} .$$

The fact that S_0 is the set of all non-positive real numbers, and that, for each positive integer m , $S_m \cap T''_m = \emptyset$, $S_m \cup T''_m = (m - 1 , m - 1/2]$, which is $[I_m \cup I'_m] \cap R' (F;m-1/2)$, and $(m - 1/2 , m) = I'_m \cap R'_c(F;\infty)$, forces us to define T to be the following unions of Borel sets in $R_w(F;\infty)$:

$$\bigcup_{i=1}^{\infty} T''_i = \bigcup_{m=1}^{\infty} T_m \ .$$

The most convenient definition of h over T is the following: if x is a point of T , and if m is that positive integer such that T_m contains x , then the value of h at x is $h(x) = y_m = 2(1-m)$. However, other definitions are possible; for example, by observing that

$$\sum_{i=5}^{11} P[T''_i] = \sum_{i=110}^{139} P[T''_i] \ ,$$

and by letting

$$M = \bigcup_{i=5}^{11} T''_i \ , \ N = \bigcup_{i=110}^{139} T''_i \ , \ T^*_1 = (T_1 \cup N) - M \ ,$$

$T^*_2 = (T_2 \cup M) - N$ and, for each integer $m > 2$, $T^*_m = T_m$, we could define h over T as follows: if x is a point of T , and if m is that positive integer such that T^*_m contains x , then $h(x) = y_m$. Nevertheless, our choice of definition of h over T has the advantage that its value at each point x of T is

$$\text{(3)} \qquad h(x) = -2\,\gamma\left(\frac{4x-1}{404}\right) .$$

Since the formulas (1) , (2) and (3) are all the same, the function h has, thus far, been defined so that it coincides with a convenient monotone non-increasing step-function over the Borel set $R - S$.

It remains for us to define h over S . Such a definition will be given at the end of a process which follows that in the proof of the theorem in Chapter II of ARL 65-75 .

First, we define the function K to be that function which maps R into the set of all Borel subsets of the set $T \cup R(F;\infty)$ in such a way that, at each point y of R, the value of K is the Borel set

$$K(y) = \{x : x \in T \cup R(F;\infty) \text{ and } h(x) \leq y\} .$$

In order to put this definition of K into a more tractable form, we define the following sets: for each positive integer m, let X_m be that subset of $R(F;\infty)$ which contains exactly 101 points, the n'th of which is $x_{m,n} = 101(m-1)+n$; let

$$X'_m = \bigcup_{i=1}^{m} X_i \text{ ; let } L_m = T_m \cup X_m \text{ ; let } L'_m = \bigcup_{i=1}^{m} L_i \text{ ; and let } L = \bigcup_{i=1}^{\infty} L_i .$$

Since $L'_m = T'_m \cup X'_m$ for each positive integer m, $L = T \cup R(F;\infty)$. Thus, if y is a point of J_0 , then $K(y) = L$; and, if, for some positive integer m, y is a point of $J_m \cup J'_m = [-2m , 2 - 2m)$, then $K(y) = L - L'_m$. Since $P[L] = 101/400$ and, for each positive integer m, $P[X_m] = z_m = 1/4m(m+1)$,

$$P[X'_m] = \sum_{i=1}^{m} P[X_i] = \frac{m}{4(m+1)} \quad \text{and} \quad P[L'_m] = P[T'_m] + P[X'_m] = \frac{101m}{400(m+1)} ,$$

it follows that, if $y \in J_0$, then $P[K(y)] = 101/400$; and, if, for some positive integer m, $y \in J_m \cup J'_m$, then $P[K(y)] = 101/400(m+1)$.

Next, we define the function g to be that function in B whose value at each point y of R is $g(y) = G(y) - P[K(y)]$. Thus, if $y \in J_0$, then $g(y) = (299y+198)/400(y+2)$; and, if, for some positive integer m , $y \in J_m \cup J'_m$, then $g(y)$ is either $(100y+299m)/400m(m+1)$ or $99/400m$

according as y is or is not less than $99/100 - 2m$.

Now, we intend to define a function V which maps the closed interval $[0, s]$, where $s = P[S] = 299/400$, into the set of all Borel subsets of S in such a way that each of the following four conditions is satisfied:

(i) for each point u of $[0, s]$, the image $V(u)$ of u under the mapping V is a Borel subset of S such that $P[V(u)] = u$;

(ii) for any points u and v of $[0, s]$ such that $u < v$, $V(u)$ is a proper subset of $V(v)$;

(iii) if $u = 0$, then $V(u) = \emptyset$; and

(iv) if $u = s$, then $V(u) = S$.

The particular function V which we shall define is taken from a non-denumerably infinite collection of functions, each of which satisfies these four conditions, but no two of which yield the same definition of h over S .

Let X denote that subset of $R - S$ which contains the point x of $R - S$ if, and only if, there exists a positive integer m such that $x = m - 101/200$; and let μ be that function which maps $S \cup X$ into the class of all Borel subsets of S in such a way that the image of each point $x \in S \cup X$ under the mapping μ is the set $\mu(x) = S \cap R'(F;x)$ if $x \in S$, and is the set $\mu(x) = S'_m$ if x is the point $m - 101/200$ of X . Thus, for each point x of $S \cup X$, either $x \leq 0$ and $P[\mu(x)] = F(x)$, so that $P[\mu(x)] = 101/2(101-x)$, or there exists a positive integer m such that $m-1 < x < m$ and $P[\mu(x)] = q'_m - q_m + 2w_m(x-m+1) = q'_{m-1} + 2w_m(x-m+1)$, so that $P[\mu(x)] = (20200x+299m^2+29900m+2030201)/400(m+100)(m+101)$. If x and z are points of $S \cup X$ such that $x < z$, then

$$0 = P[\emptyset] < P[\mu(x)] < P[\mu(z)] < P[S] = s ,$$

so that $\mu(x)$ is a proper subset of $\mu(z)$. Consequently, the function μ is a one-to-one mapping of Dom μ , its domain of definition $S \cup X$, onto Ran μ , its range of values, which is a proper subset of the set of all proper Borel subsets of S ; in other words, μ defines a one-to-one correspondence between the elements (points) of Dom μ and the elements (sets) of Ran μ . Furthermore, there is a one-to-one correspondence between the points of the open interval $(0 , s)$ and the points of the set $S \cup X =$ Dom μ such that $u \in (0, s)$ and $x \in S \cup X$ correspond to one another if, and only if, $P[\mu(x)] = u$. Therefore, there is a one-to-one correspondence between the points of $(0, s)$ and the Borel sets in Ran μ such that the point $u \in (0 , s)$ and the set $Z \in$ Ran μ correspond to one another if, and only if, $P[Z] = u$.

We define the function V , whose domain of definition Dom V is the closed interval $[0 , s]$, in such a way that its range of values Ran V is Ran μ augmented by the values $V(0) = \emptyset$ and $V(s) = S$ of V at 0 and s , respectively, and its value at each point u of $(0 , s) =$ Ran μ is that set $V(u) = \mu(x)$ in Ran μ , for some unique point x of Dom μ , which is such that $P[\mu(x)] = u$. Thus, if $u = 0$, then $V(u) = \emptyset$; if $0 < u \leq \frac{1}{2}$, then $V(u) = \mu\left(\frac{101(2u-1)}{2u}\right)$; if, for some positive integer m , $q'_{m-1} < u \leq q'_m$, then $V(u) = \mu\left(\frac{400(m+100)(m+101)u-299m^2-29900m-2030201}{20200}\right)$; and, if $u = s = 299/400$, then $V(u) = S$.

Let C be that function which maps R onto $\text{Ran}\ \mu$ in such a way that its value at each point y of R is $C(y) = V(g(y))$. Thus, if, for some positive integer m , $-2m \leq y < \frac{99}{100} - 2m$, then

$C(y) = \mu\left(\frac{101(100y-200m^2+99m)}{100y+299m}\right)$; if, for some positive integer m ,

$\frac{99}{100} - 2m \leq y < 2 - 2m$, then $C(y) = \mu\left(\frac{101(99-200m)}{99}\right)$; if $0 \leq y \leq \frac{202}{99}$,

then $C(y) = \mu\left(\frac{101(99y-202)}{299y+198}\right)$; and, if, for some positive integer m ,

$\frac{400m+20002}{9999} < y \leq \frac{400m+20402}{9999}$, then $C(y) = \mu\left(\frac{299m+9799}{200} - \frac{2(m^2+201m+10100)}{101(y+2)}\right)$.

Let C' be that function which maps R into the set of all subsets of R in such a way that its value at each point y of R is that subset $C'(y)$ of $C(y)$ which contains the point x of $C(y)$ if, and only if, for each real number $z < y$, x is not a point of $C(z)$. Thus, if, for some positive integer m , $-2m \leq y < 2 - 2m$, then $C'(y)$ is either the empty set $\emptyset$ or the set whose only point is $101(100y-200m^2+99m)/(100y+299m)$ according as y does or does not exceed $99/100 - 2m$; if $0 \leq y \leq 202/99$, then $C'(y)$ is the set whose only point is $101(99y-202)/(299y+198)$; and, if, for some positive integer m , $(400m+20002)/9999 < y \leq (400m+20402)/9999$, then $C'(y)$ is the set whose only point is $\frac{299m+9799}{200} - \frac{2(m^2+201m+10100)}{101(y+2)}$.

Finally, we define the function h over the set S in such a way that its value at each point x of S is that point $h(x) = y$ of R which is such that $C'(y) = x$. Thus, if, for some positive integer m ,

$-\frac{101(200m+101)}{99} \leq x < -\frac{101(200m-99)}{99}$, then $h(x) = \frac{m(299x+20200m-9999)}{100(101-x)}$; if

$-\frac{10201}{99} \leq x \leq 0$, then $h(x) = \frac{2(99x+10201)}{9999-299x}$; and, if, for some positive integer m , $m - 1 < x < m - \frac{101}{200}$, then $h(x) = \frac{400(101x+m^2+50m+5151)+2m+202}{101(299m+9799-200x)}$.

There can be little doubt that the method used in this construction of a function h in (f,*,G) leaves much to be desired. After having shown that (F,G) is a compatible pair, we might have found an easier way of defining a function in (F,*,G) . However, it does not seem likely that there is a procedure which is necessarily always most convenient for dealing with problems of this kind. In fact, this particular illustration involves very few of the difficulties which are anticipated in the theory of Chapter II of ARL 65-75 .

Of course, as was mentioned earlier in this section, it is possible that a decision on the matter of whether or not a given ordered pair (F,G) of functions in D is a compatible pair might be precluded by some inherent difficulties in the definitions of F and G . This possibility arises when one takes a mathematical intuitionist's critical point of view and observes the shaky foundations under the measure-theoretic notions which were entertained in the proof of the theorem of Chapter II of ARL 65-75 .

Now, once it has been shown by a satisfactory procedure that a given ordered pair (F,G) in $D \times D$ is a compatible pair, and that a definite function h is in the non-empty set (F,*,G) , the matter of defining other functions than h in (F,*,G) can be considered. The next section is devoted to this consideration.

6. Some Members of the Class C_2

The class C_2 is that set of subsets of B which contains the subset M of B if, and only if, there exist functions F and G in D such that (F,*,G) = M . In this section, we want to do two things: —— we would like to show how the theory in Chapter II of ARL 65-75 suggests a way of generating special subsets of the non-empty members of the class C_2 by employing the group of all one-one, Lebesgue-measure-preserving mappings of the closed interval $I = [0,1]$ onto itself; —— and we would like to provide illustrative material of sufficient complexity to justify a few of the elaborate details of the theory which was constructed around class C_2 in ARL 65-75. In order to realize the latter intention, we shall make use of some pathological functions throughout the discussion; and we shall define these functions by means of sprays. The subject "Spray-forms and Sprays" is considered in section 3 of the Appendix in ARL 65-75; however, the brief discussion of this subject which follows should be adequate for our present needs.

A sequence (by which term we mean an infinite sequence) may be thought of as an array of things, called its terms, which have been entered one after another into the places of a figurative structure which we shall call the sequence-form. If the entries (i.e., the terms) of a particular sequence are real numbers, then the sequence is an element of the set W which was defined in section 2. However, a particular sequence, which is produced by entering into each place of the sequence-form a sequence of real numbers, one, and only one, of which is non-zero and is 1 , is a matrix in the set T which was defined in section 2. By such examples as the latter, one could justify the idea that matrices are special sequences. For this reason, we do not hesitate to bypass matrices in our search for a generalization of the sequential notion.

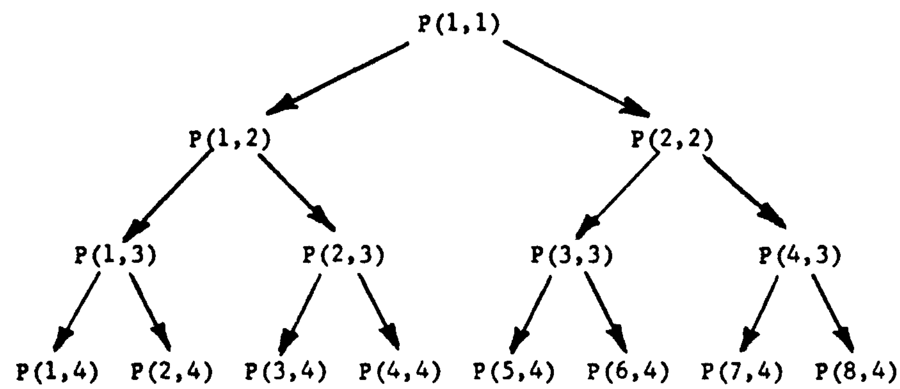
For each natural number n, a term in the fundamental sequence of natural numbers, let $P(1,n)$ denote one, and only one, place in the sequence-form; and, for each place in the sequence-form, let there be one, and only one, natural number n such that $P(1,n)$ denotes that place. Each place $P(1,n)$ in the sequence-form has one, and only one, successor: the place $P(1,n+1)$; each place $P(1,n)$ which is different from $P(1,1)$ has one, and only one, predecessor: the place $P(1,n-1)$; and the place $P(1,1)$ has no predecessor.

By analogy, we produce a generalized version of the sequence-form which, for a positive integer k, is called the k-footed spray-form. For each ordered pair of positive integers (m,n), in which $m \leq k^{n-1}$, let $P(m,n)$ denote one, and only one, place in the k-footed spray-form; and, for each place in the k-footed spray-form, let there be one, and only one, ordered pair of positive integers (m,n), in which $m \leq k^{n-1}$, such that $P(m,n)$ denotes that place. Each place $P(m,n)$ in the k-footed spray-form has k, and only k, successors, the t'th of which is the place $P(km-k+t,n+1)$; each place $P(m,n)$ which is different from $P(1,1)$ has one, and only one, predecessor; the place $P(h,n-1)$, in which h is the greatest integer less than $(m+k)/k$; and the place $P(1,1)$ has no predecessor.

The one-footed spray-form is the sequence form. The two-footed spray-form, which we shall call simply the spray-form, will now be given our attention. Figure 1 is an illustrative array of fifteen of the symbols which denote places of the spray-form; in it, the fourteen arrows are directed from seven of the place symbols toward the symbols which denote their successors.

Figure 1

The Spray-Form



We shall refer to the first and second successors of a place as its *left* and *right feet*, respectively; and we shall refer to the predecessor of a place as its *head*. The place $P(m,n)$ is said to be in the m'th *position* on the n'th *level* of the spray-form. The place $P(1,1)$ is called the *topmost place*. A sequence of places of the spray-form, in which the place $P(m,n)$ is the first term and each term, a place, is followed immediately by either its left foot or its right foot, is said to be a *limb* of $P(m,n)$. A limb, all of whose terms after the first term are left (or right) feet, is called a *left limb* (or a *right limb*); and a limb which is either a left limb or a right limb is called a *straight limb*. Thus, each place of the spray-form has exactly two straight limbs: a right limb and a left limb.

The left limb of a place $P(m,n)$ which is not a term of the left limb of P(1,1) has as its so-called heart that place whose right foot's left limb has P(m,n) as one of its terms; the right limb of a place P(m,n) which is not a term of the right limb of P(1,1) has as its heart that place whose left foot's right limb has P(m,n) as one of its terms; the left limb of a place which is a term of the left limb of P(1,1) has no heart; and the right limb of a place which is a term of the right limb of P(1,1) has no heart. The left and right limbs of P(4,4) have the respective hearts P(2,3) and P(1,1) ; the left and right limbs of P(81,8) have the respective hearts P(3,3) and P(41,7) ; the right limb of P(1,3) has the heart P(1,2) ; and the left limb of P(1,3) has no heart.

By the branch of a place P(m,n) is meant that part of the spray-form which contains a place P(a,b) if, and only if, P(a,b) is a term of a limb of P(m,n) . A branch is structurally the same as the spray-form.

A proper limb of the spray-form is any limb of the topmost place which is not the left limb of P(1,1) , and which has infinitely many terms that are left feet.

A spray (for "spreading array") is envisioned as being the array which is produced by entering objects, called entries, into the places of the spray-form; this is analogous to the treatment of a sequence as the result of having entered objects, called terms, into the places of the sequence-form.

By a limb L' in a spray S (which is not the same as a limb of a place in the spray-form) is meant the sequence of entries in the terms of a proper limb L of the spray-form; the limb L' in S is said to occupy the proper limb L of the spray-form.

We shall have need to consider only those sprays whose entries are real numbers, and whose limbs are convergent sequences. Each such spray S is called a valuable spray; and the limit of each limb in S is said to be the value of the limb. If S is a valuable spray, then S is said to spread over that subset M of the real line R which contains the real number x if, and only if, there is a limb in S whose value is x . If S is a valuable spray, and if no two limbs in S have the same value, then S is said to be a real spray.

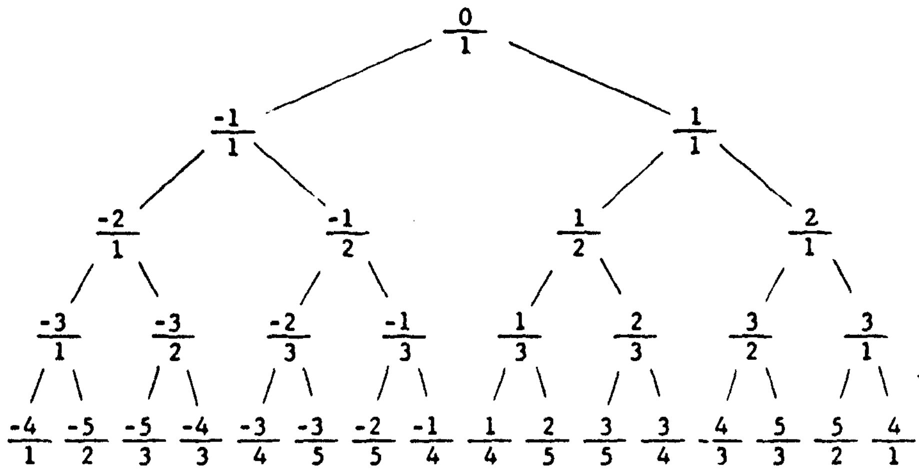
If S is a real spray which spreads over R , and if T is a valuable spray, then the symbol $S \rightarrow T$ denotes that function which maps R into R in such a way that, at each point x of R , its value y in R is the value of that limb in T which occupies the same proper limb of the spray-form that is occupied by the limb in S whose value is x .

We are now ready to define five special valuable sprays. This is done for the purpose of facilitating our definition of some pathological Baire functions in B .

The spray X is constructed in such a way that each of its entries is a ratio of two integers whose greatest common divisor is 1 . For each positive integer n , its entries in the places $P(1,n)$ and $P(2^{n-1},n)$ of the spray-form are $\frac{1-n}{1}$ and $\frac{n-1}{1}$, respectively. The sequence of numerators of its entries in the places of a straight limb of the spray-form which has a heart is an arithmetic progression whose common difference is the numerator of its entry in that heart; and the sequence of denominators of its entries in the places of a straight limb which has a heart is an arithmetic progression whose common difference is the denominator of its entry in that heart. The first five levels of spray X are shown in Figure 2.

Figure 2

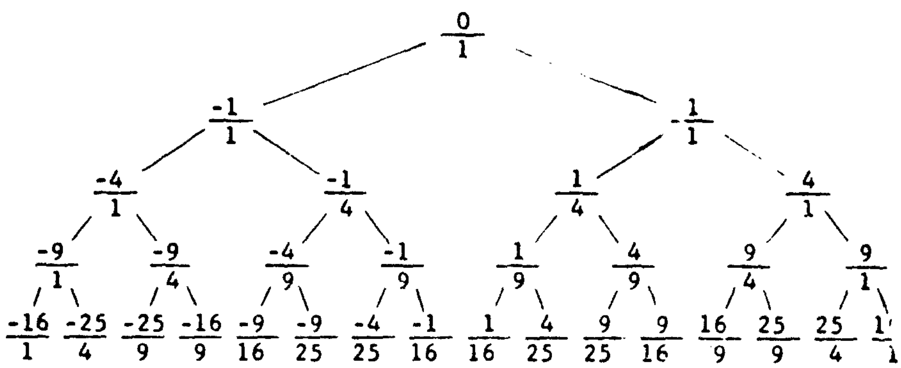
Spray X



The spray $\underline{X}$ is constructed in such a way that, for any positive integers m and n which are such that $2^{n-2} < m \leq 2^{n-1}$, its entry in $P(m,n)$ is the square of the entry of spray X in $P(m,n)$; and, for any integers m and n which are such that $1 \leq m \leq 2^{n-1}$, its entry in $P(m,n)$ is the negative of its entry in $P(2^{n-1}+1-m,n)$. The first five levels of spray $\underline{X}$ are shown in Figure 3.

Figure 3

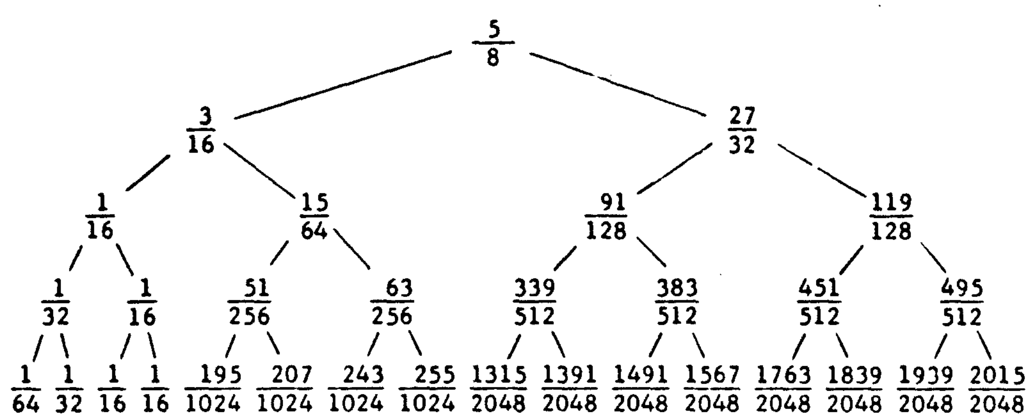
Spray $\underline{X}$



The spray Y is constructed in such a way that each of its entries is a ratio of two integers whose greatest common divisor is 1. Its entries in P(1,1) , P(1,2) , P(2,2) and P(2,3) are $\frac{5}{8}$, $\frac{3}{16}$, $\frac{27}{32}$ and $\frac{15}{64}$, respectively. For each integer $n > 2$, its entry in P(1,n) is $1/2^{n+1}$; and its entry in each place of the branch of P(2,n+1) is $1/2^{n+1}$. If $\frac{a}{b}$ is its entry in any place of the branch of P(2,3) , then its entries in the left and right feet of the place occupied by $\frac{a}{b}$ are $\frac{4a-9}{4b}$ and $\frac{4a+3}{4b}$, respectively; and, if $\frac{a}{b}$ is its entry in any place of the branch of P(2,2) , then its entries in the left and right feet of the place occupied by $\frac{a}{b}$ are $\frac{4a-9-\sqrt{2b}}{4b}$ and $\frac{4a+3+\sqrt{2b}}{4b}$, respectively. The first five levels of Y are shown in Figure 4.

Figure 4

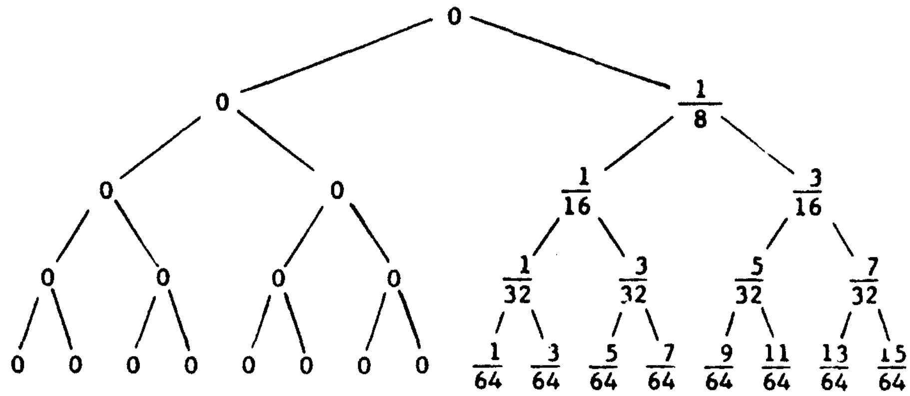
Spray Y



The spray Z is constructed in such a way that its entry in each place of the branch of P(2,2) is a ratio of two integers whose greatest common divisor is 1; and its entry in each place which does not belong to the branch of P(2,2) is 0. Its entry in P(2,2) is $\frac{1}{8}$; and, if $\frac{a}{b}$ is its entry in any place of the branch of P(2,2) , then its entries in the left and right feet of this place are $\frac{2a-1}{2b}$ and $\frac{2a+1}{2b}$, respectively. The first five levels of Z are shown in Figure 5.

Figure 5

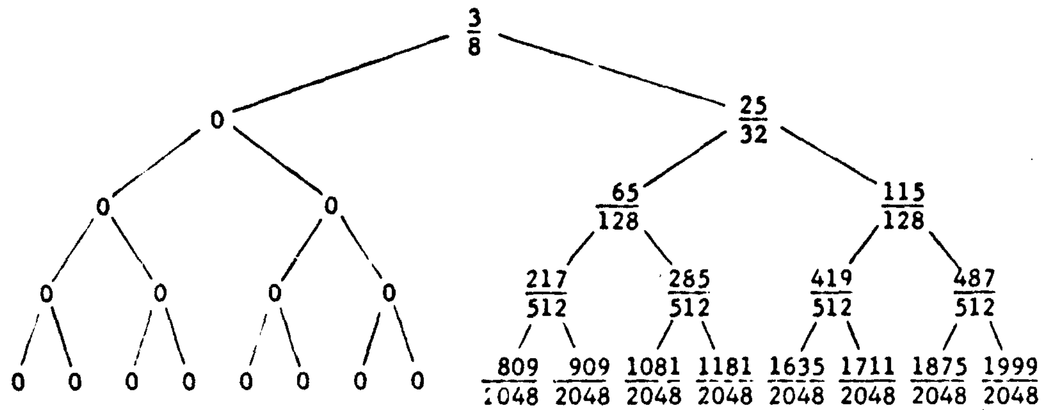
Spray Z



The spray V is constructed in such a way that its entry in each place of the branch of P(1,2) is 0 ; and its entry in each place which does not belong to the branch of P(1,2) is a ratio of two integers whose greatest common divisor is 1 . Its entries in P(1,1) , P(2,2) and P(3,3) are $\frac{3}{8}$, $\frac{25}{32}$ and $\frac{65}{128}$, respectively. If $\frac{a}{b}$ is its entry in any place of the branch of P(3,3) , then its entries in the left and right feet of this place are $\frac{4a-27-\sqrt{2b}}{4b}$ and $\frac{4a+9+\sqrt{2b}}{4b}$, respectively. For any integer $n > 2$, the entries of V in $P(2^{n-1},n)$ and $P(2^n-1,n+1)$ are $(2^{2n+1}-3 \cdot 2^{n-1}-1)/2^{2n+1}$ and $(2^{2n+3}-5 \cdot 2^{n+1}-13)/2^{2n+3}$, respectively; and, if $\frac{a}{b}$ is its entry in any place of the branch of $P(2^n-1,n+1)$, then its entries in the left and right feet of this place are $\frac{4a-9-\sqrt{2b}}{4b}$ and $\frac{4a+3+\sqrt{2b}}{4b}$, respectively. The first five levels of spray V are shown in Figure 6.

Figure 6

Spray V



By making use of the sprays X , Y , Z , $\underline{X}$ and V , the Baire functions $\underline{F}$, $\tilde{\underline{F}}$ and $\underline{G}$ in B are defined as follows: $\underline{F} = X \longrightarrow Y$, $\tilde{\underline{F}} = X \longrightarrow Z$ and $\underline{G} = \underline{X} \longrightarrow V$. Now that the letters X , Y , Z and V have served their purpose in the definition of the special functions $\underline{F}$, $\tilde{\underline{F}}$ and $\underline{G}$, we free them for other uses in this discussion.

Some particular values of these special functions are: $\underline{F}(-3/2) = 1/16$, $\underline{F}(-101/100) = 1/16$, $\underline{F}(-1) = 3/16$, $\underline{F}(-2/3) = 51/256$, $\underline{F}((1-\sqrt{5})/2) = 1/5$, $\underline{F}(0) = 5/8$, $\underline{F}(\sqrt{2}) = 74/85$, $\underline{F}(3/2) = 451/512$, $\tilde{\underline{F}}(0) = 0$, $\tilde{\underline{F}}(1) = 1/8$, $\tilde{\underline{F}}(\sqrt{2}) = 3/20$, $\tilde{\underline{F}}(3/2) = 5/32$, $\underline{G}(-1/100) = 0$, $\underline{G}(0) = 3/8$, $\underline{G}(4/25) = 909/2048$, $\underline{G}(9/16) = 1181/2048$, $\underline{G}(1) = 25/32$ and $\underline{G}(3/2) = 3439/4368$.

Throughout the remainder of this section, we shall have occasion to repeat some of the definitions given previously. This will be done in the process of providing the kind of illustrative material which we hope will support the procedure that was followed in the proof of the theorem in Chapter II of ARL 65-75.

The symbol B_1 denotes the class of all Borel sets in R. The symbol B denotes the set of all Baire functions which map R into R. The symbol I denotes the closed interval $[0,1]$ in R. The symbol D denotes the set of all distribution functions which map R into I. For example, the special functions $\underline{F}$, $\bar{\underline{F}}$ and $\underline{G}$ are in B; and the functions $\underline{F}$ and $\underline{G}$ are in D.

For any $h \in B$, the symbol $\bar{h}$ denotes that function which maps R into B_1 in such a way that its value at each $y \in R$ is the Borel set $\bar{h}(y) = \{x : h(x) \leq y\}$. The special symbol ω denotes that function in B whose value at each $x \in R$ is $\omega(x) = x$; therefore, for each $x \in R$, $\bar{\omega}(x)$ denotes the half-line $(-\infty, x]$.

For any $F \in D$, the symbol P_F denotes that probability measure which maps B_1 into I in such a way that, for each $x \in R$, the P_F-measure of the Borel set $\bar{\omega}(x)$ is $P_F[\bar{\omega}(x)] = F(x)$. For any $F \in D$ and $h \in B$, there exists one, and only one, function G in D whose value at each $y \in R$ is $G(y) = P_F[\bar{h}(y)]$; and the fact that the functions F, h and G are related to one another in this way is stated symbolically by " (F,h,G) ". For any $F \in D$ and $G \in D$, the symbol $(F,*,G)$ denotes that subset of B which contains $h \in B$ if, and only if, (F,h,G) is a true statement. For example, the special set $(\underline{F},*,\underline{G})$ contains the special function $\underline{h}$ in B whose value at each $x \in R$ is $\underline{h}(x) = x^2$. We intend to describe a method of obtaining various functions in a set $(F,*,G)$ when one such function is known.

For any $F \in D$, the symbol $R(F;\infty)$ denotes the set of all points of discontinuity of F; and the symbol $R'(F;\infty)$ denotes the difference set $R - R(F;\infty)$. For example, $R(\underline{F};\infty)$ contains $x \in R$ if, and only if, x is either a negative integer or a rational number in the open half-line $(-1,\infty)$; and $R(\underline{G};\infty)$ contains $y \in R$ if, and only if, y is the square of a non-negative rational number.

For any $F \in D$, the symbol $R_c(F;\infty)$ denotes that subset of $R'(F;\infty)$ which contains $x \in R'(F;\infty)$ if, and only if, for each real number $r < x$, $F(r) < F(x)$; and the symbol $R_c'(F;\infty)$ denotes the difference set $R'(F;\infty) - R_c(F;\infty)$. Furthermore, for each $x \in R$, the symbol $R_c(F;x)$ denotes the intersection set $R_c(F;\infty) \cap \bar{\omega}(x)$; and the symbol $R_c(F;-\infty)$ denotes the empty set $\emptyset$. If $R_c'(F;\infty) \neq \emptyset$, then it is either a single interval with no left endpoint or the union of pairwise disjoint intervals with no left endpoints; and, over each such interval in the composition of $R_c'(F;\infty)$, the function F has only one value (i.e., it has a so-called "plateau"). For example, $R_c'(\underline{F};\infty)$ is the set of all non-integers of the half-line $\bar{\omega}(-1)$. Despite the fact that the arbitrary value $F(x)$ of a function $F \in D$ increases as x increases over $R_c(F;\infty)$, it need not be true that, if X is a Borel subset of $R_c(F;\infty)$ whose Lebesgue measure is positive, then $P_F[X]$ is positive. It happens, for example, that, if X is that subset of $R_c(\underline{F};\infty)$ which consists of all the irrational numbers in the open interval $(-1,0)$, then $P_{\underline{F}}[X] = 0$.

Let $\bar{R}$ denote that special set which consists entirely of all the points of R and the two special non-real points ∞ and $-\infty$. For any $F \in D$, let the symbol $\bar{F}$ denote the function which maps $\bar{R}$ into I in such a way that its value at each point x of $\bar{R}$ is $P_F[R_c(F;x)]$. For example, if F is $\underline{F}$, then, for each $x \in R$, $\bar{F}(x) = \bar{\underline{F}}(x)$.

For any $F \in D$, the symbol $R_w(F;\infty)$ denotes that subset of $R_c(F;\infty)$

which contains $x \in R_c(F;\infty)$ if, and only if, for each real number $r < x$, $\overline{F}(r) < \overline{F}(x)$. For example, $R_w(\underline{F};\infty)$ is the set of all positive irrational numbers; and the set $R_w(\underline{G};\infty)$ is the set of all positive real numbers which are not the squares of rational numbers.

For any $F \in D$, let the symbol I_F denote the closed interval $[0,\overline{F}(\infty)]$ onto which the set $\overline{R}$ is mapped by the function $\overline{F}$; let the symbol $\hat{F}$ denote that function which maps $R_w(F;\infty)$ into I_F in such a way that its value at each $x \in R_w(F;\infty)$ is $\hat{F}(x) = \overline{F}(x)$; let the symbol H_F denote that subset of I_F which contains $u \in I_F$ if, and only if, there exists an $x \in R_w(F;\infty)$ at which the value of $\hat{F}$ is $\hat{F}(x) = u$; and let the symbol J_F denote that subset of I_F which is such that $J_F \cup H_F = I_F$ and $J_F \cap H_F = \emptyset$. Since F is a one-one mapping of $R_w(F;\infty)$ onto H_F, it has a unique inverse which may be denoted by the symbol $\hat{F}^{-1}$. Since the Lebesgue measure of J_F is 0, the Lebesgue measure of H_F is the length $\overline{F}(\infty)$ of the interval I_F. For example, since $I_{\underline{F}}$ is the closed interval $[0,1/4]$, and since $J_{\underline{F}}$ is that subset of $I_{\underline{F}}$ which contains 0 and a positive real number u if, and only if, there exist integers m and n which are such that $0 < m \leq 2^{n-2} + \frac{1}{2}$ and $(2m-1)/2^{n+1} = u$, it follows that $J_{\underline{F}}$ contains a denumerably infinite number of points, so that its Lebesgue measure is 0; consequently, the Lebesgue measure of both the set $H_{\underline{F}} = I_{\underline{F}} - J_{\underline{F}}$ and the interval $I_{\underline{F}}$ is the latter's length $1/4$.

If α is a one-one mapping of a set S onto itself, and if X is a subset of S, then, by the <u>modified</u> <u>restriction</u> of α to the subset X of S, we mean the unique one-one mapping β of X onto itself which is defined as follows: let $Y = S - X$; let $X^{(0)}$ be that subset of Y which contains $y \in Y$ if, and only if, $\alpha^{-1}(y) = x$ is an element of X, where α^{-1} denotes the unique inverse of α; for each positive integer n, let $X^{(n)}$ be that subset of X which contains $x \in X$ if, and only if, $\alpha(x)$ is an element of $X^{(n-1)}$; let X' be that subset of X which contains $x \in X$ if, and only if, there exists a positive integer n such that $x \in X^{(n)}$; let $X'' = X - X'$; and let β be that function over X whose value at each $x \in X$ is either $\beta(x) = x$ or $\beta(x) = \alpha(x)$ according as $x \in X'$ or $x \in X''$. Since, for each non-negative integer n, and for each $x \in X^{(n)}$, $\alpha^{-1}(x)$ is either an element of Y or an element of the subset $X^{(n+1)}$ of X, there can be no more elements in $X^{(n+1)}$ than there are in $X^{(n)}$; consequently, if S is a Borel subset of R, if α is a one-one, Lebesgue-measure-preserving mapping of S onto S, if X is a Borel subset of S, and if $Y = S - X$ is a set with Lebesgue measure 0, then every Borel subset of X' has Lebesgue measure 0.

Let M denote the set of all one-one, Lebesgue-measure-preserving mappings of the closed interval I onto itself. For any $\alpha \in M$ and $\beta \in M$, let the symbol $\beta\alpha$ denote that function in M whose value at each $u \in I$ is $\beta\alpha(u) = \beta(\alpha(u))$; and let the symbol α^{-1} denote the inverse of α. The set M constitutes a group with respect to the binary operation that is indicated when two symbols which denote functions in M are placed side by side in order to form a composite symbol which denotes a function in M.

For any $\alpha \in M$ and $F \in D$, let the symbol $\bar{\alpha}_F$ denote that one-one, Lebesgue-measure-preserving mapping of the closed interval I_F onto I_F whose value at each $u \in I_F$ is either $\bar{\alpha}_F(u) = u$ or $\bar{\alpha}_F(u) = \bar{F}(\infty)\alpha(u/\bar{F}(\infty))$ according as $\bar{F}(\infty)$ is or is not 0; and let the symbol α_F denote the modified restriction of $\bar{\alpha}_F$ to the subset H_F of I_F.

Let T denote that function which maps $M \times D \times B$ into B in such a way that, for each ordered triple (α, F, h) in $M \times D \times B$, the image in B of (α, F, h) is the function $T_\alpha F{:}h$ whose value at each $x \in R$ is either $T_\alpha F{:}h(x) = h(\hat{F}^{-1}(\alpha_F(\hat{F}(x))))$ or $T_\alpha F{:}h(x) = h(x)$ according as x is or is not a point of $R_w(F;\infty)$. The image in B of $(\alpha, F, h) \in M \times D \times B$ under the mapping T is denoted by the symbol $T_\alpha F{:}h$ in order that one may be able to denote some other functions by convenient symbols. Thus, for any $\alpha \in M$, the symbol T_α denotes that function which maps $D \times B$ into B in such a way that, under T_α, the image in B of the ordered pair $(F,h) \in D \times B$ is $T_\alpha F{:}h$; and, for any ordered pair (α, F) in $M \times D$, the symbol (or operator) $T_\alpha F{:}$ denotes that function which maps B into B in such a way that, under $T_\alpha F{:}$, the image in B of the function $h \in B$ is $T_\alpha F{:}h$. If (α, F, h) is any ordered triple in $M \times D \times B$, and if G is that function in D which makes (F,h,G) a true statement, then the subset $(F,*,G)$ of B contains both h and $T_\alpha F{:}h$.

For each ordered triple (a,b,c) in $I \times I \times I$ which is such that $a \leq b$ and $b-a \leq c$, let the symbol $(a{:}b{:}c)$ denote that function in M whose value at each $u \in I$ is either $(a{:}b{:}c)(u) = a+c-u$ or $(a{:}b{:}c)(u) = u$ according as u is or is not a point of the union of closed intervals $[a,b] \cup [a-b+c,c]$.

For our illustration of this method of using an operator $T_\alpha \underline{F}{:}$ in order to obtain a function in $(\underline{F},*,\underline{G})$ which differs from $\underline{h} \in (\underline{F},*,\underline{G})$ over a subset of R with positive $P_{\underline{F}}$-measure, we choose α to be the convenient mapping $(1/4{:}1/3{:}3/4)$ in M. Then, since $\underline{h}$ is that function in B whose value at each $x \in R$ is $\underline{h}(x) = x^2$, the function $\underline{k} = T_\alpha \underline{F}{:}\underline{h}$

in B is such that its value at each $x \in R$ is either $\underline{k}(x) = 1/x^2$ or $\underline{k}(x) = x^2$ according as x is or is not an irrational number in the union of closed intervals $[1/2,(\sqrt{5}-1)/2] \cup [(\sqrt{5}+1)/2,2]$.

III. ROBUSTNESS OF TESTS

1. An Illustration of the Notion of Robustness of a Test

In order to put a particular notion of robustness of tests into a somewhat realistic and tractable, hypothetical, experimental situation, we present the following illustration:

The officer in charge of a military radar operators' school wishes to consider a change in the training regimen that has been followed for several years to one which is much less costly. A group of n entrants is put into an experimental program for the full duration of the course of training. It has been the practice at this school to assign to each trainee a proficiency rating at the completion of his training. A recent study of the school's records seems to justify the assumption that the proficiency ratings are normally distributed. Therefore, next to each such rating in the school's records, but not in the service records of the graduates, is placed a standardized rating, which is that transformation of the given proficiency rating that adjusts the aggregate of all ratings to conform with the assumption that they are normally distributed with mean 0 and variance 1. It is intended that the proficiency ratings of the trainees in the experimental group shall be transformed (i.e., shall be standardized) by the same function that is used on all the proficiency ratings. And, in order to meet the requirements of the Analysis of Variance methodology in testing the worth of the proposed change in the course, it is assumed that, if for each positive integer $i \leq n$, X_i is the random variable whose values are the possible standardized ratings that the i'th trainee in the experimental course might achieve upon the completion of his training, then the variance of X_i is 1 . Two further reasonable assumptions are made: if $i \neq j$, then X_i and X_j are independent random variables; and, for each positive integer $i \leq n$, X_i is normally distributed with mean μ . The distribution function and the probability density function of X_i , which is $N(\mu,1)$, are denoted simply by Φ_μ and ϕ_μ , respectively; and Φ_0 and ϕ_0 may be denoted simply by Φ and ϕ , respectively. Furthermore, the probability measure which induces Φ is denoted by P . Thus, the sample mean of the standardized ratings of the experimental units is the normal random variable $\bar{X} = \frac{X_1 + \cdots + X_n}{n}$ with mean μ and variance $\frac{1}{n}$.

Now, a real number α in the interval (0,1) is chosen; and, for each real number v , the symbol $c(v,\alpha)$ denotes that real number which satisfies the following equation: $P\{\bar{X} < c(v,\alpha) : \mu = v\} = \alpha$. The number $c(v,\alpha)$ is the right endpoint of an open half-line critical region Γ .

With all these things under consideration, the question of whether or not the regular training program should be replaced by the experimental training program is taken up in the following test which tests $H_0 : \mu = 0$ against the $H_A : \mu < 0$.

If $\bar{X} < c(0,\alpha)$, then reject the null hypothesis H_0 ;

and if $\bar{X} \geq c(0,\alpha)$, then do not reject H_0 .

The power function of this test is defined at each real point v by

$$\beta(v) = P\{\bar{X} < c(0,\alpha) : \mu = v\} = P\left\{\frac{\bar{X}-v}{1/\sqrt{n}} < \frac{c(0,\alpha)-v}{1/\sqrt{n}}\right\} = \Phi\left(\frac{c(0,\alpha)-v}{1/\sqrt{n}}\right).$$

Since $\alpha = P\{\bar{X} < c(0,\alpha) : \mu = 0\} = P\{\bar{X}\sqrt{n} < c(0,\alpha)\sqrt{n}\} = \Phi(c(0,\alpha)\sqrt{n})$, the real number $\gamma = c(0,\alpha)$ is $\Phi^{-1}(\alpha)/\sqrt{n}$, so that, at each real point v , the value of β is $\beta(v) = \Phi(\Phi^{-1}(\alpha) - v\sqrt{n})$.

Now, suppose that the underlying assumption that the variance of X_i is 1 is changed; that is, suppose that, for each positive integer $i \leq n$, X_i is $N(\mu,\sigma^2)$. Under this assumption, $\bar{X}$ is $N(\mu,\sigma^2/n)$. Then too, the power function β_σ of the above test, in which the critical region Γ is the open half-line with right endpoint γ , has as its value at each real point v the number

$$\beta_\sigma(v) = P\{\bar{X} < \gamma : \mu = v\} = P\left\{ \frac{\bar{X}-v}{\sigma/\sqrt{n}} < \frac{\gamma-v}{\sigma/\sqrt{n}} \right\}$$

$$= \Phi\left(\frac{\gamma-v}{\sigma/\sqrt{n}} \right) = \Phi\left(\frac{\Phi^{-1}(\alpha)-v\sqrt{n}}{\sigma} \right) .$$

For convenience, the functions δ and Δ are defined at each real point v as follows: $\delta(v) = \beta_\sigma(v) - \beta(v)$ and $\Delta(v) = |\delta(v)|$. The first derivative of δ is δ' whose value at each real point v is

$$\delta'(v) = -\frac{\sqrt{n}}{\sigma}\,\phi\left(\frac{\Phi^{-1}(\alpha)-v\sqrt{n}}{\sigma} \right) + \sqrt{n}\;\phi(\Phi^{-1}(\alpha) - v\sqrt{n}) .$$

For $\sigma^2 \neq 1$, $\delta'(v) = 0$ if, and only if, v is either

$$\frac{1}{\sqrt{n}}\left[\Phi^{-1}(\alpha) - \sigma\sqrt{\frac{2\ln\sigma}{\sigma^2-1}} \right] \quad \text{or} \quad \frac{1}{\sqrt{n}}\left[\Phi^{-1}(\alpha) + \sigma\sqrt{\frac{2\ln\sigma}{\sigma^2-1}} \right] .$$

Therefore, for each real number v

$$0 \leq \Delta(v) \leq \left| \Phi\left(\sqrt{\frac{2\ln\sigma}{\sigma^2-1}} \right) - \Phi\left(\sigma\sqrt{\frac{2\ln\sigma}{\sigma^2-1}} \right) \right| \leq \sqrt{\frac{(\sigma-1)\ln\sigma}{\pi(\sigma+1)}} ,$$

where the upper bound on the right is useful only if σ is close to 1 . In a rough way, it can be said that, with respect to the above mentioned change in the underlying assumptions, the measure of robustness of the test under discussion is

$$1 - \left| \Phi\left(\sqrt{\frac{2 \ln \sigma}{\sigma^2 - 1}}\right) - \Phi\left(\sigma\sqrt{\frac{2 \ln \sigma}{\sigma^2 - 1}}\right) \right| .$$

This measure is greater than 1/2 (i.e., the test is more than 50% robust) if σ is in the interval (e^{-1}, e) .

2. A Generalization of the Notion of Robustness of a Test

In the last section, we tried to illustrate a way of measuring the robustness of a test. By considering a contrived experimental situation and a convenient test, we produced a number between 0 and 1 which was to serve as a rough indicator of the ineffectiveness of a particular change in the assumptions underlying the test as a disturbing influence on its power function. This number, which could be expressed as a percentage, was called the measure of robustness of the test with respect to the particular change in its underlying assumptions.

In this section, we shall give a skeletal version of a generalization of the illustration in the last section. In doing this, we begin with a review of some notational conventions which, in somewhat more detail, are treated in ARL 65-75.

For each positive integer n , let R_n denote the Euclidean space of n dimensions; for each point x of R_n and each positive integer $t \leq n$, let x_t denote the t'th coordinate of x ; let $R = R_1$;

let B_n denote the class of all Borel subsets of R_n ; for each point x of R_n , let $[-\infty,x]$ denote that member of B_n which contains the point u of R_n if, and only if, for each positive integer $t \leq n$, $u_t \leq x_t$; let D_n denote the set of all distribution functions which map R_n into the closed interval $I = [0,1]$ in R ; let M_n denote the set of all probability measures which map B_n into I ; for each $F \in D_n$, let $Pr(F)$ denote that probability measure P in M_n which induces F , so that, for each point x of R_n , $P(u : u \in [-\infty,x]) = F(x)$; and, for each $P \in M_n$, let $Df(P)$ denote that distribution function F in D_n which is induced by P , so that $Pr(F) = P$. Furthermore, for each ordered pair of positive integers (m,n) , let ${}_mB_n$ denote the set of all Baire functions which map R_n into R_m ; and, for each ordered pair (F,h) in the Cartesian product set $D_n \times {}_mB_n$, let $(F,h,*)$ denote that subset of D_m which contains the function G in D_m if, and only if, for each point y of R_m , $G(y) = P[x : h(x) \in [-\infty,y]]$, where $P = Pr(F)$.

If, for positive integers i and j , V denotes a non-empty subset of R_i and W denotes a non-empty subset of R_j , then the symbol $V:W$ denotes that subset of R_{i+j} which contains the point u of R_{i+j} if, and only if, there exist points v of V and w of W such that, for each positive integer $t \leq i + j$, u_t is either w_{t-i} or v_t according as t does or does not exceed i ; and, if this point u is in $V:W$ because such points v of V and w of W do exist, then u may be denoted alternatively by the symbol $v:w$. If $V \subset R$ and $W \subset R$, then $V:W$ is the Cartesian product set $V \times W$; however, if $V \subset R$ and $W \subset R_2$, then $V:W$ need not be a Cartesian product subset of R_3 .

If S is any non-empty set, then a function ρ which maps the Cartesian product set $S \times S$ into R is said to be a metric for S if, and only if, the following three conditions are satisfied:

(1) if, for any ordered pair (a,b) of $S \times S$, $\rho(a,b) = 0$, then $a = b$;

(2) for each $a \in S$, $\rho(a,a) = 0$; and

(3) for each ordered triple (a,b,c) of $S \times S \times S$, $\rho(a,c) + \rho(b,c) \geq \rho(a,b)$.

Consequently, for each $(a,b) \in S \times S$, $\rho(a,b) = [\rho(a,b) + \rho(a,b)]/2 \geq \rho(a,a)/2$, which, by (2), is zero, so that $\rho(a,b)$ is non-negative; and, furthermore, $\rho(a,b) = \rho(b,b) + \rho(a,b)$ by (2), $\rho(b,b) + \rho(a,b) \geq \rho(b,a)$ by (3), $\rho(b,a) = \rho(a,a) + \rho(b,a)$ by (2) and $\rho(a,a) + \rho(b,a) \geq \rho(a,b)$ by (3), so that $\rho(a,b) \geq \rho(b,a) \geq \rho(a,b)$ and $\rho(a,b) = \rho(b,a)$. If the range of values of ρ is a subset of $I = [0,1]$ in R , then ρ is said to be a metric for S limited to I .

New, let i , j , m , and n be positive integers; let the non-empty subsets V and W of R_i and R_j be called parameter spaces; let the member Γ of B_m be called a critical region; and let the function h of ${}_mB_n$ be called a transformation. For each point v:w of V:W , let $F_{v:w}$ be a unique distribution function in D_n ; let $P_{v:w} = Pr(F_{v:w})$; and let $G_{v:w}$ be that distribution function in the subset $(F_{v:w},h,*)$ of D_m . For each point w of W , let β_w , called a power function, be that function which maps V into the interval I whose value at each point v of V is $\beta_w(v) = P_{v:w}(x : h(x) \in \Gamma)$. Let S be that set of functions which map V into I , which contains the function f if, and only if, there exists a $w \in W$ such that $\beta_w = f$. Finally, let ρ be a metric for S limited to I . For example, $\rho(f,g)$ could be the least upper bound of $|f(v) - g(v)|$, where $f \in S$, $g \in S$ and $v \in V$.

Suppose that, in an experimental situation, a decision is to be based on the way a random variable X whose range of values is R_n is distributed. It is thought that the distribution function of X is $F_{r:s}$, where $r:s \in V:W$. A test of the hypothesis that r is the proper parameter in V is designed; and it makes use of the transformation h as well as the critical region Γ which is such that $\beta_s(r)$ is some small number α in I . This test may be stated as follows: if $h(X) \in \Gamma$, reject H_0 ; otherwise, do not reject H_0 . The power function of this test is β_s , where the parameter s is a fixed point of W and, hence, is an underlying assumption of the test. In ascribing to this test some measure of robustness with respect to a change in the underlying assumption concerning the parameter s , another test, with h and Γ unchanged, but with s changed to $w \in W$ yields the power function β_w . And the number $1 - \rho(\beta_w,\beta_s)$ serves as a measure of the robustness of the test with respect to the particular change in its underlying assumptions.

Of course, the utility of such a measure of robustness of a test will depend on the choice and general acceptance of the metric ρ . Even though these measures of robustness are in I and are intuitively appealing, it is not easy to defend the choice of any one of them with general considerations.

3. The Kolmogorov Metric on the Space of Location Families of Distribution Functions

In order to define the notions of robustness of a statistical test and robustness of a transformation, some quantitative measure of the distance between power functions and the distance between distribution functions appears to be highly desirable. It would appear that several authors (see, e.g., [1]-[10]) have held this view to a greater or lesser extent. In section 1, an example was given which exploits the Kolmogorov metric on the space of normal distribution functions as a measure of the robustness of a test. The power functions being dealt with in this example are very simply related to distribution functions, a situation which will not occur in general. However, for those situations for which it does occur, it appears that the assumption of normality does not play a large part in determining the outcome. The present section is an attempt to investigate the Kolmogorov metric on the space of distribution functions of location-parameter families.

We begin with some definitions and notation. Let I be a subset of the real line, and let $\{F(x;\mu) : \mu \in I\}$ be a family of distribution functions on the real line. We may suppose that the distribution functions are right-continuous, and that the ordering on I is that induced by the natural order on the real line. We will call the family $\{F(x;\mu) : \mu \in I\}$ a location-parameter family with respect to I if and only if there is a distribution function G such that for each real x, $F(x;\mu) = G(x-\mu)$, for each $\mu \in I$. It is obvious that such a family is stochastically increasing; i.e., if $\lambda \in I$ and $\mu \in I$ with $\lambda > \mu$, then $F(x;\lambda) \leq F(x;\mu)$, for all real x.

Since we will require a measure of distance between distribution functions, we introduce one such measure, often called the Kolmogorov metric. For each pair μ, λ with $\mu \in I$ and $\lambda \in I$, and for each pair F_μ, F_λ of distribution functions, we set $d(x;\mu,\lambda) = |F(x;\mu) - F(x;\lambda)|$ and $\rho(F_\mu,F_\lambda) = \sup_x d(x;\mu,\lambda)$. It is readily verified that ρ is indeed a metric on the space of distribution functions. As noted above, a location-parameter family (with respect to I) is stochastically ordered. In view of this, we may dispense with the absolute value signs in the definition of $d(x;\mu,\lambda)$ if we assume that $\lambda > \mu$.

Because of the exploratory nature of our study, we will make a series of increasingly stringent assumptions on the location-parameter family $\{F(x;\mu) : \mu \in I\}$, and at each stage, study the effect of each additional assumption on the Kolmogorov metric.

Assumption 1. G is a continuous function.

Consequence. $d(x;\mu,\lambda)$ is a continuous function of x.

Proof. $d(x;\mu,\lambda) = F(x;\mu) - F(x;\lambda) = G(x-\mu) - G(x-\lambda)$.

Assumption 2. G is a strictly increasing function.

Consequence. To each positive real number η, there correspond real numbers $a(\eta)$ and $b(\eta)$, with $a(\eta) \leq b(\eta)$, such that $x < a(\eta)$ implies $d(x;\mu,\lambda) < \eta$ and $b(\eta) < x$ implies $d(x;\mu,\lambda) < \eta$.

Proof. From $\lambda > \mu$ and G strictly increasing it follows that $F(x;\mu) = G(x-\mu) > G(x-\lambda) = F(x;\lambda)$, for all real x. The remainder of the proof follows from the fact that G is a distribution function.

We may summarize the consequences of assumptions 1 and 2 as follows:

__Remark__. If G is a continuous strictly increasing function, then $d(\cdot;\mu,\lambda)$ is a continuous function which attains its (absolute) maximum on a finite interval.

Proof. $d(\cdot;\mu,\lambda)$ is, when restricted to $[a(\eta),b(\eta)]$, a continuous function on a closed, bounded interval, and hence it attains its (absolute) maximum on $[a(\eta),b(\eta)]$.

Assumption 3. $G(x) = 1 - G(-x)$, for all real x; i.e., G is a symmetric distribution function.

Consequence. $d(\cdot;\mu,\lambda)$ is symmetric about the point $\frac{1}{2}(\mu+\lambda)$.

Proof. $d(\frac{1}{2}(\mu+\lambda) - x;\mu,\lambda) = G(\frac{1}{2}(\mu+\lambda)-x-\mu) - G(\frac{1}{2}(\mu+\lambda)-x-\lambda)$

$= G(\frac{1}{2}(\lambda-\mu)-x) - G(\frac{1}{2}(\mu-\lambda)-x) = 1 - G(x - \frac{1}{2}(\lambda-\mu)) + G(x - \frac{1}{2}(\mu-\lambda)) - 1$

$= G(x + \frac{1}{2}(\lambda-\mu)) - G(x + \frac{1}{2}(\mu-\lambda)) = G(\frac{1}{2}(\mu+\lambda) + x-\mu) - G(\frac{1}{2}(\mu+\lambda)+ x-\lambda)$

$= d(\frac{1}{2}(\mu+\lambda) + x;\mu,\lambda)$, for every real x.

Assumption 4. G is absolutely continuous, with probability density function g.

Consequence. $d'(\cdot;\mu,\lambda)$ satisfies $d'(\frac{1}{2}(\mu+\lambda) + x;\mu,\lambda) = -d'(\frac{1}{2}(\mu+\lambda) - x;\mu,\lambda)$, for all real x.

Proof. For each real number x, define $k(x;\mu,\lambda)$ by $k(x;\mu,\lambda) = d(\frac{1}{2}(\mu+\lambda) + x;\mu,\lambda)$. From the consequence of assumption 3, it follows that $k(\cdot;\mu,\lambda)$ is an even function. Since the derivative (when it exists) of an even function is an odd function, the proof is complete.

We note that assumptions 3 and 4 together yield the following result.

__Remark__. g is an even function.

Proof. The existence of g follows from assumption 4 and its

...follows from assumption 3.

Assumption 5. g is increasing for $x < 0$.

Consequence. g is decreasing for $x > 0$, and hence g has a single maximum at $x = 0$.

Proof. This is a simple consequence of the assumption and the fact that g is an even function.

Assumption 6. g has a derivative g' ; thus, $g' \equiv G''$.

Consequence. If $x > 0$, then $g'(x) < 0$.

Proof. Since g is decreasing for $x > 0$, it follows that its derivative. g' , must be negative for $x > 0$.

We can now summarize the consequences of these assumptions in the following way.

<u>Theorem</u>. If G is an absolutely continuous function which is symmetric about zero, and if G'' exists and is an increasing function for $x < 0$, then $\rho(F_\mu, F_\lambda) = 2G(\frac{1}{2}(\lambda-\mu)) - 1$, for $\lambda > \mu$.

Proof. Because $\lambda > \mu$, $\rho(F_\mu, F_\lambda) = \sup_x [F(x;\mu) - F(x;\lambda)]$. It is clear that $d(\frac{1}{2}(\mu+\lambda);\mu,\lambda) = 2G(\frac{1}{2}(\lambda-\mu)) - 1$, and that $\sup_x [F(x;\mu) - F(x;\lambda)]$ $= d(\frac{1}{2}(\mu+\lambda);\mu,\lambda)$ follows from these facts: (1) $d'(\frac{1}{2}(\mu+\lambda);\mu,\lambda) =$ $g(\frac{1}{2}(\lambda-\mu)) - g(\frac{1}{2}(\mu-\lambda)) = 0$, and (2) $d''(\frac{1}{2}(\mu+\lambda);\mu,\lambda) = g'(\frac{1}{2}(\lambda-\mu)) - g'(\frac{1}{2}(\mu-\lambda))$ $= 2g'(\frac{1}{2}(\lambda-\mu)) < 0$.

As we will show presently, this theorem furnishes sufficient, but not necessary, conditions under which the distance, as measured by the Kolmogorov metric, between two members of a particular family of distribution functions can be computed as a function of the distance between their indices. We list here some cases of interest that are included in the above analysis.

1. Normal; $I = R_1$; $F(x;\mu) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} \exp\,[-\frac{1}{2}(t-\mu)^2]dt$.

2. Double exponential; $I = R_1$; $F(x;\mu) = \frac{1}{2} \int_{-\infty}^{x} \exp\,[-|t-\mu|]dt$.

3. Cauchy; $I = R_1$; $F(x;\mu) = \begin{cases} \frac{1}{\pi} \arctan\,[1+(x-\mu)^2] \ , & x < \mu \\ \frac{1}{2} + \frac{1}{\pi} \arctan\,[1+(x-\mu)^2] \ , & \mu \le x \end{cases}$.

4. Student's t ; $I = R_1$;

$$F_n(x;\mu) = \Gamma\left(\frac{n+1}{2}\right)\left[\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}\right]^{-1} \int_{-\infty}^{x} [1+n^{-1}(t-\mu)^2]^{-\frac{n+1}{2}}\, dt \ ,$$

for each natural number n .

5. Gamma; $I = R_1$; $F_{\alpha,\beta}(x;\mu) = \beta^{-\alpha}[\Gamma(\alpha)]^{-1} \int_{\mu}^{x} (t-\mu)^{\alpha-1} \exp\,[-\beta^{-1}(t-\mu)]dt$,

for fixed real positive α and fixed real positive β , $x > \mu$.

We turn now to a location-parameter family of distribution functions with nonsymmetric members. Take $I = R_1$, and for each $\mu \in I$, let

$$F(x;\mu) = \begin{cases} 0 & , \ x \le \mu \\ 1 - \exp(\mu-x) & , \ \mu < x \end{cases} \quad . \tag{1}$$

It is clear that the family $\{F(x;\mu) : \mu \in I\}$ is a location-parameter family with respect to I . However, the distribution function G for which $F(x;\mu) = G(x-\mu)$ for each real x is

$$G(y) = \begin{cases} 0 & , \ y \le 0 \\ 1 - \exp(-y) & , \ 0 < y \end{cases} \quad , \tag{2}$$

which is not symmetric with respect to any real number. Nevertheless, we may still prove the following result.

<u>Theorem</u>. If $F(x;\mu)$ and $G(y)$ are given by (1) and (2) respectively, then for any $\mu \in I$ and $\lambda \in I$, with $\lambda > \mu$, we have $\rho(F_\mu, F_\lambda) = G(\lambda-\mu)$.

Proof. We distinguish three distinct and exhaustive cases: (i) $-\infty < x < \mu$, (ii) $\mu \leq x \leq \lambda$, and (iii) $\lambda < x < \infty$.

(i) For $x < \mu$, $F(x;\mu) = F(x;\lambda) = 0$; thus $d(x;\mu,\lambda) = 0$, for $x < \mu$.

(ii) For $\mu \leq x \leq \lambda$, $d(x;\mu,\lambda)$ can be written as $d(x;\mu,\lambda) = F(x;\mu) - 0 = 1 - \exp(\mu-x)$, which exhibits the fact that for $\mu \leq x \leq \lambda$, $d(x;\mu,\lambda)$ is a strictly monotone increasing, continuous and bounded function of x. Hence $d(x;\mu,\lambda)$ attains its maximum value (which is clearly $G(\lambda-\mu)$) at $x = \lambda$.

(iii) We complete the proof by considering this case, since for $\lambda < x$, $d(x;\mu,\lambda) = e^{-x}(e^{\lambda}-e^{\mu}) < e^{-\lambda}(e^{\lambda}-e^{\mu}) = 1 - \exp(\lambda-\mu) = G(\lambda-\mu)$.

Let us attempt to generalize this theorem to include location-parameter families whose members are not necessarily symmetric. Thus, we suppose that for some subset I of R_1, we have a family $\{F(x;\mu) : \mu \in I\}$ of distribution functions such that for each $\mu \in I$, $F(x;\mu) = 0$, for $x < \mu$. If we further suppose that the family $\{F(x;\mu) : \mu \in I\}$ is a location-parameter family with respect to I, with $F(x;\mu) = G(x-\mu)$ for all real x, then, as before, this family is stochastically increasing. But we do not assume now that G is symmetric with respect to any real number. Without adding any restriction, we suppose that $\lambda > \mu$, and we observe that for $-\infty < x < \mu$, we have

$F(x;\lambda) = F(x;\mu) = 0$, so that for $x < \mu$, $d(x;\mu,\lambda) = 0$.

If we now suppose that G is a continuous function which is strictly increasing whenever it is positive, then as above, corresponding to each positive real number η , there is a real number $a(\eta)$ such that $a(\eta) < x$ implies $d(x;\mu,\lambda) < \eta$. When $d(x;\mu,\lambda)$ is restricted to the closed interval $[\mu,a(\eta)]$, it is a continuous function on a closed interval (which is also bounded), and hence it attains its (absolute) maximum on $[\mu,a(\eta)]$. From the nature of $d(x;\mu,\lambda)$ for $x \in [\mu,\lambda]$, we may conclude that $\sup_{x\in[\mu,\lambda]} d(x;\mu,\lambda) = \sup_{x\in[\mu,\lambda]} G(x-\mu) = G(\lambda-\mu)$. While it may not always be true that $\rho(F_\mu,F_\lambda) = G(\lambda-\mu)$ for the situation under consideration, we are able to state the following result.

<u>Theorem</u>. Suppose G is a continuous function which is strictly increasing for positive values of its argument. If, for each positive real number δ less than $a(\eta)-\lambda$, $G(\lambda-\mu+\delta) < G(\lambda-\mu) + G(\delta)$, then $\rho(F_\mu,F_\lambda) = G(\lambda-\mu)$.

Proof. If $x \in (\lambda,a(\eta)]$, then $(x-\lambda) \in (0,a(\eta)-\lambda]$, so that $x-\lambda$ may be taken as δ . If this substitution is made in the inequality on G , then we see that $G(x-\mu) < G(\lambda-\mu) + G(x-\lambda)$, which is equivalent to $d(x;\mu,\lambda) < G(\lambda-\mu)$ for $x \in (\lambda,a(\eta)]$, thus completing the proof of the theorm.

To see that the exponential example treated in equations (1) and (2) is indeed covered by the inequality of the theorem, it is sufficient to note that the distribution function of the exponential distribution given by (2) satisfies the functional equation

$$G(x+y) - G(y) = G(x)[1-G(y)]$$

for all $x \geq 0$ and $y \geq 0$. Since $\lambda > \mu$ and $\delta > 0$, we see that for the G of equation (2) we have

$$G(\lambda-\mu+\delta) - G(\delta) = G(\lambda-\mu)[1-G(\delta)] ,$$

and hence that the inequality of the theorem is satisfied, since for any $\delta > 0$, we have $G(\delta) > 0$.

REFERENCES

[1] Box, G. E. P. and Cox, D. R. (1964), An analysis of transformations, J.R.S.S., Series B, 26, 211-252.

[2] Box, G. E. P. and Tiao, G. C. (1962), A further look at robustness via Bayes's theorem, Biometrika, 49, 419-432.

[3] Box, G. E. P. and Tiao, G. C. (1964), A Bayesian approach to the importance of assumptions applied to the comparison of variances, Biometrika, 51, 153-167.

[4] Box, G. E. P. and Tiao, G. C. (1965), Multiparameter problems from a Bayesian point of view, Ann. Math. Statist., 36, [illegible]

[5] Lawton, W. H. (1965), Some inequalities for central and non-central distributions, Ann. Math. Statist., 36, 1521-1525.

[6] Pfanzagl, J. (1964), On the topological structure of some ordered families of distributions, Ann. Math. Statist., 35, 1216-1228.

[7] Severo, N. C. and Olds, E. G. (1956), A comparison of tests on the mean of a logarithmico-normal distribution with known variance, Ann. Math. Statist., 27, 670-686.

[8] Severo, N. C. (1957), Asymptotic behavior of tests on the mean of a logarithmico-normal distribution with known variance, Ann. Math. Statist., 28, 1044-1046.

[9] Tukey. J. W. (1957), On the comparative anatomy of transformations, Ann. Math. Statist., 28, 602-632.

[10] Tukey, J. W. (1958), A problem of Berkson, and minimum variance orderly estimators, Ann. Math. Statist., 29, 588-592.

IV. TWO THEOREMS ON SOLUTIONS OF DIFFERENTIAL-DIFFERENCE EQUATIONS AND APPLICATIONS TO EPIDEMIC THEORY

1. Introduction

We present two theorems that provide simple iterative solutions of special systems of differential-difference equations. We show as examples of the theorems the simple stochastic epidemic (cf. Bailey, 1957, p. 39, and Bailey, 1963) and the general stochastic epidemic (cf. Bailey, 1957; Gani, 1965; and Siskind, 1965), in each of which we let the initial distribution of the number of uninfected susceptibles and the number of infectives be arbitrary but assume the total population size bounded. In all of the references cited above the methods of solution involve solving a corresponding partial differential equation, whereas we deal directly with the original system of ordinary differential-difference equations. Furthermore in the cited references the authors begin at time $t = 0$ with a population having a fixed number of uninfected susceptibles and a fixed number of infectives. For the simple stochastic epidemic with arbitrary initial distribution we provide solutions not obtainable by the results given by Bailey (1957 or 1963). For the general stochastic epidemic, if we use the results of Gani or Siskind, then the solution of the problem having an arbitrary initial distribution would involve additional steps that would sum proportionally-weighted conditional results.

Let $\underset{\sim}{x}(t)$ and $\underset{\sim}{x}'(t)$ denote n by 1 column matrices whose i'th row elements are the real-valued differentiable function $x_i(t)$ and its derivative $x_i'(t)$, respectively, each defined for all $t \geq 0$. Let initial conditions at time $t = 0$ be $\underset{\sim}{x}(0) = \underset{\sim}{a}$, where the column matrix $\underset{\sim}{a}$ has as i'th row element the real number a_i. Let $\underset{\sim}{B}$ be an n by n triangular matrix whose (i,j)'th element b(i,j) is a constant for each pair (i,j) and in particular $b(i,j) = 0$ for $i < j$. We let $I_n = \{1,\ldots,n\}$ and for convenience of notation we denote b(i,i) by b_i for $i \in I_n$. Let $\underset{\sim}{C}$ denote an n by n matrix with c(i,j) as (i,j)'th element. Occasionally it will be convenient to write b(i,j) or c(i,j) as b[i,j] or c[i,j], respectively. The n by 1 column matrix with i'th row element $\exp(b_i t)$ is denoted by $\underset{\sim}{e}(t)$. We define the symbol $\delta_1(x)$ to be equal to t when $x = 0$ and x^{-1} when $x \neq 0$. We define the real-valued function $\delta_2(x)$ to be equal to 1 when $x = 0$ and x^{-1} when $x \neq 0$. Finally, we shall make frequent use of the function $\mathcal{E}(x)$ defined as 1 for $x \geq 0$ and 0 for $x < 0$.

2. Solutions of Some Systems of Differential-Difference Equations

THEOREM 1. Let $\underset{\sim}{x}(0) = \underset{\sim}{a}$, and for $t \geq 0$ let $\underset{\sim}{x}'(t) = \underset{\sim}{B}\underset{\sim}{x}(t)$, where (i) $b(i,j) = 0$ for $i-j \geq 2$, and (ii) for each $i \in I_n$, $b_i = b_j$ for at most one $j \in I_n$ and $j \neq i$. Then $\underset{\sim}{x}(t) = \underset{\sim}{C}\underset{\sim}{e}(t)$, where

$$c(i,j) = \begin{cases} 0 \quad , i < j \\ a_1 \quad , i = j = 1 \\ b(i,i-1)[c_1(i-1,j)\delta_1(b_j-b_i) - c_2(i-1,j)\delta_1^2(b_j-b_i) \\ \qquad\qquad + c_2(i-1,j)\delta_1(b_j-b_i)t] \quad , i > j \\ a_i - \sum_{u=1}^{i-1} c_1(i,u) \quad , i = j > 1 \quad , \end{cases} \tag{1}$$

in which for fixed j the functions c_1 and c_2 are defined recursively (in i) as the term independent of t and the coefficient of t, respectively, in $c(i,j)$; i.e.,

$$c(i,j) = c_1(i,j) + c_2(i,j)t \quad . \tag{2}$$

(Thus, in particular, $c_1(1,1) = a_1$, $c_2(1,1) = 0$; and for $i > 1$,

$$c_1(i,i) = a_i - \sum_{u=1}^{i-1} c_1(i,u) \quad , \quad c_2(i,i) = 0 \quad .)$$

<u>Proof.</u> Note that $x_1(t) = a_1 e^{b_1 t}$ so that $c(1,j)$ is equal to a_1 when $j = 1$ and 0 when $j > 1$. For fixed integer i_0 where $1 \leq i_0 - 1 < n$ assume equation (1) holds for positive integers $i \leq i_0 - 1$. Consider the equation

$$x'_{i_0}(t) - b_{i_0} x_{i_0}(t) = b(i_0, i_0 - 1) \sum_{j=1}^{i_0-1} c(i_0 - 1, j) e^{b_j t} \quad . \tag{3}$$

The term in which $j = j_0 \leq i_0 - 1$ on the right hand side of equation (1) contributes to the solution of $x_{i_0}(t)$ the term $K(i_0, j_0) e^{b_{j_0} t}$. We shall show that $K(i_0, j_0) = c(i_0, j_0)$. Note that

$$K(i_0,j_0)e^{b_{j_0}t} = e^{b_{i_0}t}\, b(i_0,i_0-1)\Big[c_1(i_0-1,j_0)\int e^{(b_{j_0} - b_{i_0})t}\,dt \tag{4}$$
$$+ c_2(i_0-1,j_0)\int t e^{(b_{j_0} - b_{i_0})t}\,dt\Big] \; .$$

Our proof distinguishes three cases: Case 1. $b_{j_0} \neq b_i$ for

$i = j_0 + 1, \ldots, i_0$. For this case $c_2(i_0-1, j_0) = 0$. Therefore $K(i_0, j_0) = b(i_0, i_0-1)c_1(i_0-1, j_0)(b_{j_0} - b_{i_0})^{-1}$. Case 2. $b_{i_0} = b_{j_0}$. Here again $c_2(i_0-1, j_0) = 0$, but now $K(i_0, j_0) = b(i_0, i_0-1)c_1(i_0-1, j_0)t$. Case 3. $b_{j_0} = b_{k_0}$ where $j_0 < k_0 < i_0$. Then

$$K(i_0, j_0) = b(i_0, i_0-1)[c_1(i_0-1, j_0)(b_{j_0} - b_{i_0})^{-1} - c_2(i_0-1, j_0)(b_{j_0} - b_{i_0})^{-2} + c_2(i_0-1, j_0)(b_{j_0} - b_{i_0})^{-1}t] .$$

Note that all three cases are accounted for by the δ_1 symbol as used in equation (1) where $i_0 > j_0$. Therefore

$$x_{i_0}(t) = \sum_{j=1}^{i_0-1} c(i_0, j)e^{b_j t} + K_1 e^{b_{i_0} t} ,$$

and so by applying the initial condition $x_{i_0}(0) = a_{i_0}$ we obtain $K_1 = a_{i_0} - \sum_{u=1}^{i_0-1} c_1(i_0, u)$. This completes the proof of Theorem 1.

THEOREM 2. If $\underset{\sim}{x}(0) = \underset{\sim}{a}$, $\underset{\sim}{x}'(t) = \underset{\sim}{B}\underset{\sim}{x}(t)$ for $t \geq 0$, and if for every pair of integers $\alpha < \beta$ such that $b_\alpha = b_\beta$ we have either (i) $b(\alpha + \gamma, \alpha) = 0$ for $\gamma = 1, \ldots, \beta - \alpha$, or (ii) $b(\beta, j) = 0$ for $j = \alpha, \ldots, \beta-1$, then $\underset{\sim}{x}(t) = \underset{\sim}{C}\underset{\sim}{e}(t)$ where

$$c(i,j) = \begin{cases} 0 , & i < j \\ a_1 , & i = j = 1 \\ \delta_2(b_j - b_i) \sum_{u=j}^{i-1} b(i,u)c(u,j) , & i > j \\ a_i - \sum_{u=1}^{i-1} c(i,u) , & i = j > 1 . \end{cases} \qquad (5)$$

Proof. Note that $x_1(t) = a_1 e^{b_1 t}$ so that $c(1,j)$ is equal to a_1 when $j = 1$ and 0 when $j > 1$. For fixed integer i_0 where $1 \le i_0 - 1 < n$, assume equation (5) holds for all positive integers $i \le i_0 - 1$ and let $R_{i_0} = \{j : b_j = b_{i_0}, 1 \le j \le i_0 - 1\}$ and $R^*_{i_0} = \{j : b_j \ne b_{i_0}, 1 \le j \le i_0 - 1\}$. Then

$$x'_{i_0}(t) - b_{i_0} x_{i_0}(t) = \sum_{u=1}^{i_0-1} b(i_0,u) x_u(t)$$

$$= \sum_{u=1}^{i_0-1} b(i_0,u) \sum_{v=1}^{u} c(u,v) e^{b_v t}$$

$$= \left(\sum_{v \in R^*_{i_0}} + \sum_{v \in R_{i_0}} \right) \sum_{u=v}^{i_0-1} b(i_0,u) c(u,v) e^{b_v t} .$$

Therefore

$$x_{i_0}(t) = \sum_{v \in R^*_{i_0}} \sum_{u=v}^{i_0-1} \frac{b(i_0,u)c(u,v)}{(b_v - b_{i_0})} e^{b_v t} + D + \left(a_{i_0} - \sum_{v \in R^*_{i_0}} \sum_{u=v}^{i_0-1} \frac{b(i_0,u)c(u,v)}{(b_v - b_{i_0})} \right) e^{b_{i_0} t} , \quad (6)$$

where

$$D = \left[\sum_{v \in R_{i_0}} \sum_{u=v}^{i_0-1} b(i_0,u) c(u,v) \right] t e^{b_{i_0} t} . \quad (7)$$

If R_{i_0} is empty, then the term D does not appear in (6). If R_{i_0} is not empty, then for fixed $j_0 \in R_{i_0}$ we have

$$\sum_{u=j_0}^{i_0-1} b(i_0,u) c(u,j_0) = 0$$

if either $b(i_0,u) = 0$ for $u = j_0,\ldots,i_0-1$, or $b(j_0+\gamma,j_0) = 0$ for $\gamma = 1,\ldots,i_0-j_0$. (The latter condition is seen to be sufficient when we use the identity for $i > j$:

$$c(i,j) = \begin{cases} \delta(b_j-b_i)b(i,j)c(j,j) , & i = j+1 \\ \delta(b_j-b_i)[b(i,j)c(j,j) + \sum_{u=j+1}^{i-1} b(i,u)c(u,j)] , & i = j+2,\ldots,n \text{ .)} \end{cases} \quad (8)$$

Therefore D is identically equal to zero.

It is now easy to see that we may write

$$x_{i_0}(t) = \sum_{v=1}^{i_0} c(i_0,v)e^{b_v t} ,$$

where

$$c(i_0,v) = \begin{cases} 0 , & v > i_0 \\ \delta_2(b_v-b_{i_0}) \sum_{u=v}^{i_0-1} b(i_0,u)c(u,v) , & v < i_0 \\ a_{i_0} - \sum_{u=1}^{i_0-1} c(i_0-u) , & v = i_0 , \end{cases}$$

and so the theorem is proved.

3. Stochastic Epidemics

By an epidemic population we shall mean a well-defined set Ω of elements (individuals) ω defined to be in Ω if and only if for some time $t \geq 0$, ω is an uninfected susceptible or an infective. For each

$\omega \in \Omega$ and for each $t \geq 0$ we define:

$$\left.\begin{aligned}
W_1(\omega,t) &= \begin{cases} 1 & \text{, if } \omega \text{ is an uninfected susceptible at time } t \\ 0 & \text{, otherwise} \end{cases} \\
W_2(\omega,t) &= \begin{cases} 1 & \text{, if } \omega \text{ is an infective at time } t \\ 0 & \text{, otherwise} \end{cases} \\
W_3(\omega,t) &= \begin{cases} 1 & \text{, if } \omega \text{ is neither an uninfected susceptible nor} \\ & \quad \text{an infective at time } t \\ 0 & \text{, otherwise.} \end{cases}
\end{aligned}\right\} \tag{9}$$

We shall assume that the number of elements in Ω is M, a finite positive integer.

Let

$$\left.\begin{aligned}
R(t) &= \sum_{\omega\in\Omega} W_1(\omega,t) \\
S(t) &= \sum_{\omega\in\Omega} W_2(\omega,t) \\
L(t) &= \sum_{\omega\in\Omega} W_3(\omega,t).
\end{aligned}\right\} \tag{10}$$

Then

$$M = R(t) + S(t) + L(t) \quad . \tag{11}$$

We shall denote the size of the epidemic population at time $t \geq 0$ by $N(t)$, which consists of the total of uninfected susceptibles and infectives at time $t \geq 0$; i.e..

$$N(t) = R(t) + S(t) \quad . \tag{12}$$

The problem we consider is to find $p_{rs}(t)$, the probability that $R(t) = r$ and $S(t) = s$, when we are given the initial distribution $\{p_{rs}(0)\}$ and information about the infinitesimal transition probabilities for an ω to move amongst the three states of being (1) an uninfected susceptible, or (2) an infective, or (3) neither an uninfected susceptible nor an infective.

4. The Simple Stochastic Epidemic

In the simple stochastic epidemic, which has been extensively investigated by Bailey (1957, 1963), there is a positive integer N such that for each $t \geq 0$ the probability is one that $N(t) = N(0) = N$. Therefore $S(t) = N - R(t)$.

When we make the usual assumptions (cf. Bailey (1957), p. 39) about the infinitesimal transition probabilities, then we obtain*

$$p'_{r,N-r}(t) = (r+1)(N-r-1)p_{r+1,N-r-1}(t) - r(N-r)p_{r,N-r}(t) \tag{13}$$

for $r = 0,1,\ldots,N$, where $p_{rs}(t) \equiv 0$ if $r < 0$ or $r > N$. We write the initial conditions for this system as

$$p_{r,N-r}(0) = a_{r,N-r} , \tag{14}$$

$r = 0,1,\ldots,N$, where each $a_{r,N-r} \geq 0$ and $\sum_{r=0}^{N} a_{r,N-r} = 1$. Thus

*Because there is no loss of generality for our purpose, we have assumed throughout sections 4 and 5 that the infection rate is equal to one.

in addition to including an arbitrary initial distribution, we have introduced for completeness the case in which $r = N$, which corresponds to an initial population completely free of infectives.

We shall now put this problem into the framework of the theory of section 2.

LEMMA 1. For each ordered pair of integers $(r,N-r)$ let

$$\left.\begin{aligned} p_{r,N-r}(t) &= \mathcal{E}(r)\mathcal{E}(N-r)x_k(t) \\ a_{r,N-r} &= a_k \quad , \end{aligned}\right\} \tag{15}$$

where $k \equiv k(r;N) = N-r+1$. (16)

Then the system of equations (13) with initial conditions (14) , where $p_{r,N-r}(t) \equiv 0$ if $r < 0$ or $r > N$, is equivalent to the system

$$x_k'(t) = (N-k+2)(k-2)\mathcal{E}(k-2)x_{k-1}(t) - (N-k+1)(k-1)x_k(t) \quad , \tag{17}$$

with initial conditions

$$x_k(0) = a_k \quad , \tag{18}$$

$k \in I_n$, where $n = N + 1$.

Proof. If we make the indicated change of variables then for $r = 0,1,\ldots,N$

$$\left.\begin{aligned} p_{r,N-r}(t) &= x_k(t) \\ p'_{r,N-r}(t) &= x_k'(t) \\ p_{r+1,N-r-1}(t) &= \mathcal{E}(k-2)x_{k-1}(t) \quad , \end{aligned}\right\} \tag{19}$$

where $k = N - r + 1$. Thus k takes values $1,2,\ldots,N + 1 \equiv n$, and so equations (17) and (18) follow.

The solution of the simple epidemic is given by

THEOREM 3. If we have the system of equations (17) with initial conditions (18), $k \in I_n$, then $\underset{\sim}{x}(t) = \underset{\sim\sim}{Ce}(t)$, where $b_i = -(N-i+1)(i-1)$ and $c(i,j)$ is given by equation (1), in which $b(i,i-1) = (N-i+2)(i-2)\varepsilon(i-2)$.

Proof. We need only note that for $i = 1,\dots,n$, $b_i = b_{N-i+2}$ and $b_i \neq b_j$ for $j \neq N-i+2$, $j \neq i$. Thus Theorem 1 applies.

Example. We illustrate the simplicity of the theory by showing the details of the example in which $N = 6$, and initial distribution

$$(a_{06}, a_{15}, a_{24}, a_{33}, a_{42}, a_{51}, a_{60}) = (0, .10, .30, .25, .15, .10, .10) .$$

If we use Lemma 1, then

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \\ x_3'(t) \\ x_4'(t) \\ x_5'(t) \\ x_6'(t) \\ x_7'(t) \end{bmatrix} = \begin{bmatrix} 0 & & & & & & \\ 0 & -5 & & & & & \\ & 5 & -8 & & & & \\ & & 8 & -9 & & & \\ & & & 9 & -8 & & \\ & & & & 8 & -5 & \\ & & & & & 5 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \\ x_5(t) \\ x_6(t) \\ x_7(t) \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \end{bmatrix} = \begin{bmatrix} .10 \\ .10 \\ .15 \\ .25 \\ .30 \\ .10 \\ 0 \end{bmatrix}$$

By applying Theorem 3 we obtain:

$c(1,1) = \frac{1}{10}$ and $c(i,1) = 0$ for $i > 1$

$$c(2,2) = \frac{1}{10}$$

$$c(3,2) = 5(\frac{1}{10})(\frac{1}{3}) = \frac{1}{6}$$

$$c(4,2) = 8(\frac{1}{6})(\frac{1}{4}) = \frac{1}{3}$$

$$c(5,2) = 9(\frac{1}{3})(\frac{1}{3}) = 1$$

$$c(6,2) = 8(1)\delta_1(0) = 8t$$

$$c(7,2) = -\ 5(8)(-\ \frac{1}{5})^2 + 5(8)(-\ \frac{1}{5})t = \frac{-8}{5} - 8t$$

$$c(3,3) = \frac{3}{20} - \frac{1}{6} = -\ \frac{1}{60}$$

$$c(4,3) = 8(-\ \frac{1}{60})(\frac{1}{1}) = -\ \frac{2}{15}$$

$$c(5,3) = 9(\frac{-2}{15})\delta_1(0) = \frac{-6}{5}\ t$$

$$c(6,3) = -\ 8(\frac{-6}{5})(\frac{-1}{3})^2 + 8(\frac{-6}{5})(\frac{-1}{3})t = \frac{16}{15} + \frac{16}{5}\ t$$

$$c(7,3) = 5(\frac{16}{15})(-\ \frac{1}{8}) - 5(\frac{16}{5})(\frac{-1}{8})^2 + 5(\frac{16}{5})(-\ \frac{1}{8})t = \frac{-11}{12} - 2t$$

$$c(4,4) = \frac{1}{4} - \frac{1}{3} + \frac{2}{15} = \frac{1}{20}$$

$$c(5,4) = 9(\frac{1}{20})(-\ 1) = -\ \frac{9}{20}$$

$$c(6,4) = 8(-\ \frac{9}{20})(-\ \frac{1}{4}) = \frac{9}{10}$$

$$c(7,4) = 5(\frac{9}{10})(-\ \frac{1}{9}) = -\ \frac{1}{2}$$

$$c(5,5) = \frac{3}{10} - 1 + \frac{9}{20} = -\ \frac{1}{4}$$

$$c(6,5) = 8(-\ \frac{1}{4})(-\ \frac{1}{3}) = \frac{2}{3}$$

$$c(7,5) = 5\left(\frac{2}{3}\right)\left(-\frac{1}{8}\right) = -\frac{5}{12}$$

$$c(6,6) = \frac{1}{10} - \frac{16}{15} - \frac{9}{10} - \frac{2}{3} = -\frac{76}{30}$$

$$c(7,6) = 5\left(-\frac{76}{30}\right)\left(-\frac{1}{5}\right) = \frac{76}{30}$$

$$c(7,7) = \frac{8}{5} + \frac{11}{12} + \frac{1}{2} + \frac{5}{12} - \frac{76}{30} = \frac{9}{10}$$

Therefore the solution is

$$\begin{bmatrix} P_{60} \\ P_{51} \\ P_{42} \\ P_{33} \\ P_{24} \\ P_{15} \\ P_{06} \end{bmatrix} \equiv \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \frac{1}{60} \begin{bmatrix} 6 & & & & & & \\ 0 & 6 & & & & & \\ 0 & 10 & -1 & & & & \\ 0 & 20 & -8 & 3 & & & \\ 0 & 60 & -72t & -27 & -15 & & \\ 0 & 480t & 64+192t & 54 & 40 & -152 & \\ 0 & -96-480t & -55-120t & -30 & -25 & 152 & 54 \end{bmatrix} \begin{bmatrix} 1 \\ e^{-5t} \\ e^{-8t} \\ e^{-9t} \\ e^{-8t} \\ e^{-5t} \\ 1 \end{bmatrix}$$

$$= \frac{1}{60} \begin{bmatrix} 6 & & & & & & \\ 0 & 6 & & & & & \\ 0 & 10 & -1 & & & & \\ 0 & 20 & -8 & 3 & & & \\ 0 & 60 & -15 & -27 & -72 & & \\ 0 & -152 & 104 & 54 & 192 & 480 & \\ 0 & 56 & -80 & -30 & 120 & -480 & 54 \end{bmatrix} \begin{bmatrix} 1 \\ e^{-5t} \\ e^{-8t} \\ e^{-9t} \\ te^{-8t} \\ te^{-5t} \\ 1 \end{bmatrix}$$

Note that of the 28 entries calculated for C only 5 had $c_2 \neq 0$.

Once the individual probabilities are known then it is an easy matter to obtain other derived quantities such as the duration of the epidemic, the mean number of infectives at time t, and the distribution of the total size of the epidemic. For example, the distribution function of time to extinction may be read directly from the above solution as $G(T) = p_{06}(T) \equiv x_7(T)$ for $T \geq 0$, and $G(T) = 0$ for $T < 0$.

5. The General Stochastic Epidemic

In the general stochastic epidemic, which has recently been investigated by Gani (1965) and Siskind (1965), the probability is one that for $t_2 \geq t_1 \geq 0$, $N(t_2) \leq N(t_1)$. We shall assume the initial population bounded in the sense that there exists a positive integer N such that the probability is one that $N(0) \leq N$. Let S_N' be the set of ordered pairs of integers $\{(r,s) : r \geq 0 , s \geq 0 , r + s \leq N\}$. If we make the usual assumptions (cf. Bailey (1957), p. 53) about the infinitesimal transition probabilities then we obtain

$$p_{rs}'(t) = (r+1)(s-1)p_{r+1,s-1}(t) - s(r+\rho)p_{rs}(t) + \rho(s+1)p_{r,s+1}(t) \quad (20)$$

where ρ is the removal rate and $p_{rs}(t) \equiv 0$ if $(r,s) \notin S_N'$. We write the initial conditions for this system as

$$p_{rs}(0) = a_{rs} , \quad (21)$$

$(r,s) \in S_N'$, where $a_{rs} \geq 0$ and $\sum_{(r,s)\in S_N'} a_{rs} = 1$.

In order to put this epidemic problem into the framework of the theory of section 2 we construct a counting mechanism for the equations and the variables. This is done in the following lemma whose proof is left to the reader.

LEMMA 2. For each non-negative integer N, let S_N denote the set of ordered triplets of integers (k,r,s), where $k \geq 0$, $r \geq 0$, $s \geq 0$, $r + s \leq N$, and

$$k \equiv k(r,s;N) = (N+1)(N+2)/2 - (N+1)r - s + (r-1)r/2 \quad . \tag{22}$$

Then S_N contains exactly $n = (N+1)(N+2)/2$ ordered triplets and for each positive integer $k \leq n$, there exists one and only one ordered pair of non-negative integers (r,s) such that $(k,r,s) \in S_N$.

Therefore for each pair $(r,s) \in S_N'$ one can find $k(r,s;N)$. It might be worthwhile to point out that the converse problem can also be neatly treated; namely, for each positive integer $k \leq n$, let u be the greatest integer which is less than $(1 + \sqrt{8k+1})/2$. Then $r = N + 1 - u$ and $s = u(u+1)/2 - k$. (Later in the statement and proof of Theorem 4 it will be convenient to use the notation (r_k, s_k) in order to indicate the one-to-one correspondence between $k \equiv k(r,s;N)$ and (r_k, s_k).)

We are now prepared to effect a change of notation, which we do in the following lemma.

LEMMA 3. For each ordered pair of integers (r,s) let

$$\left.\begin{aligned} p_{rs}(t) &= \mathcal{E}(r)\mathcal{E}(s)\mathcal{E}(N-r-s)x_k(t) \\ a_{rs} &= a_k \quad , \end{aligned}\right\} \tag{23}$$

where $k \equiv k(r,s;N)$ is given by (22). Then the system of equations (20) with initial conditions (21), where $p_{rs}(t) \equiv 0$ if $(r,s) \notin S'_N$, is equivalent to the system

$$x'_k(t) + s(r+\rho)x_k(t) = (r+1)(s-1)\mathcal{E}(s-1)x_{k-N+r}(t) + \rho(s+1)\mathcal{E}(N-r-1-s)x_{k-1}(t) \quad (24)$$

with initial conditions

$$x_k(0) = a_k \quad , \quad (25)$$

$k \in I_n$, where $n = (N+1)(N+2)/2$ and $(k,r,s) \in S_N$.

Proof. If we make the indicated change of variables then for each $(k,r,s) \in S_N$ we obtain

$$\left.\begin{aligned} p_{rs}(t) &= x_k(t) \\ p'_{rs}(t) &= x'_k(t) \\ p_{r+1,s-1}(t) &= \mathcal{E}(s-1)x_{k-N+r}(t) \\ p_{r,s+1}(t) &= \mathcal{E}(N-r-1-s)x_{k-1}(t) \end{aligned}\right\} \quad (26)$$

and so the conclusion of the lemma follows immediately.

We now state the solution of the general stochastic epidemic in

THEOREM 4. If we have the system of equations (24) with initial conditions (25), $k \in I_n$, where $n = (N+1)(N+2)/2$ and $(k,r,s) \in S_N$, and if ρ is such that for $(r,s) \in S'_N$, $(r',s') \in S'_N$, $s \neq 0$ and $s' \neq 0$, we have $s(r+\rho) = s'(r'+\rho)$ only if $s = s'$, then $\underset{\sim}{x}(t) = \underset{\sim}{C}\underset{\sim}{e}(t)$, where for any $\gamma \in I_n$, $b_\gamma = -s_\gamma(r_\gamma + \rho)$, and

$$c(k,j) = \begin{cases} 0 \quad , \quad k < j \\ a_1 \quad , \quad k = j = 1 \\ [(r+1)(s-1)\mathcal{E}(s-1)c(k-N+r,j)\mathcal{E}(k-N+r-j) \\ \qquad + \rho(s+1)\mathcal{E}(N-r-1-s)c(k-1,j)]\delta_2(b_j - b_k) \quad , \quad k > j \\ a_j - \sum_{u=1}^{j-1} c(j,u) \quad , \quad k = j > 1 \ . \end{cases} \tag{27}$$

<u>Proof</u>. If $(r,s) \in S_N'$ and $s > 0$, then $b_{k(r,s;N)} = -s(r+\rho)$ are all distinct. If $s = 0$, then for $r = 0,1,\ldots,N$ we have $b_{k(r,0;N)} = 0$. The following argument holds for each positive integer $r \leq N$. In equation (24) the only possible non-zero coefficients of $x_{k(r,0;N)}$ are $b[k',k(r,0;N)]$ and $b[k'',k(r,0;N)]$, where $k' = k(r,0;N) + 1$ and $k'' = k(r,0;N) + N - r_{k''}$. We see that $k' = k(r-1,N-r+1;N)$ and $k'' = k(r-1,1;N)$. Therefore $b[k',k(r,0;N)] = \rho[(N-r+1)+1]\mathcal{E}(N-(r-1)-1-(N-r+1))$ $\rho(N-r+2)\mathcal{E}(-1) = 0$ and $b[k'',k(r,0;N)] = [(r-1)+1](1-1)\mathcal{E}(1-1) = r(0)\mathcal{E}(0) = 0$. Thus condition (i) of Theorem 2 is satisfied. Finally, by applying equation (5) we get the conclusion of our present theorem. Note the resemblance between equation (24) and equation (27) for the case $k > j$.

The question arises as to whether or not there is anything distinctive about the choice of N, the bound on the initial total population size. The following embedding theorem answers this question in the negative.

THEOREM 5. The general stochastic epidemic with total initial population size bounded by $N_1 \leq N_2$ may be treated as one with initial total population size bounded by N_2 and initial conditions satisfying

$a_{k(r,s;N_2)} = 0$ for all non-negative integer pairs (r,s) such that $r + s > N_1$.

We leave the details of the proof to the interested reader. Essentially what must be shown is that for each non-negative integer pair (r,s) such that $r + s \leq N_1$ we have $x_{k(r,s;N_1)}(t) = x_{k(r,s;N_2)}(t)$ for all $t \geq 0$. Also for each non-negative integer pair (r,s) such that $N_1 < r + s \leq N_2$ we have $x_{k(r,s;N_2)}(t) \equiv 0$.

The special cases considered by Gani and Siskind are included in Theorem 4 by simply choosing the initial conditions appropriately. In fact if a computing program is worked out for fixed N , then any of their cases in which $R(0) = r_0$ and $S(0) = s_0$, where $r_0 + s_0 = N$, may be obtained by letting $p_{r_0 s_0}(0) = a_{r_0 s_0} = 1$ and setting all other a_{rs}'s equal to zero. Furthermore, by Theorem 5, any case in which the probability is zero that the sum of $R(0)$ and $S(0)$ exceeds $N_1 < N_2$ may be obtained from the program written for N_2 by suppressing from the program all terms involving triplets $(k,r,s) \in S_{N_2}$ such that $r + s > N_1$.

Conversely, if one wishes to use the results of Gani or Siskind to solve a general stochastic epidemic with arbitrary initial distribution then one can obtain $p_{rs}(t)$ by evaluating

$$\sum_{(r_0,s_0)\in S_N'} P[R(t) = r, S(t) = s \mid R(0) = r_0, S(0) = s_0] a_{r_0 s_0} ,$$

where $P[A|B]$ is the conditional probability of A given B. This, of course, would be considerably more difficult than if one were to use our Theorem 4 directly.

<u>Example.</u> We illustrate the simplicity of the theory by showing the details of the example in which $N = 2$, $\rho = 2$, and initial distribution $(a_{00}, a_{01}, a_{02}, a_{10}, a_{11}, a_{20}) = (0, .20, .30, .25, .15, .10)$. Then $S_N = \{(1,2,0), (2,1,1), (3,1,0), (4,0,2), (5,0,1), (6,0,0)\}$.

If we use Lemma 3 then

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \\ x_3'(t) \\ x_4'(t) \\ x_5'(t) \\ x_6'(t) \end{bmatrix} = \begin{bmatrix} 0 & & & & & \\ 0 & -3 & & & & \\ & 2 & 0 & & & \\ & 1 & & -4 & & \\ & & 0 & 4 & -2 & \\ & & & & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \\ x_5(t) \\ x_6(t) \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix} = \begin{bmatrix} .10 \\ .15 \\ .25 \\ .30 \\ .20 \\ 0 \end{bmatrix}$$

By applying equation (27) we obtain

$$c(1,1) = \frac{1}{10}$$

$$c(2,1) = (\ 0 + 0c(1,1))/3 = 0$$

$$c(3,1) = (\ 0 + 2c(2,1))\delta_2(0) = 0$$

$$c(4,1) = (1c(2,1) + \ 0\ )/4 = 0$$

$$c(5,1) = (0c(3,1) + 4c(4,1))/2 = 0$$

$$c(6,1) = (\ 0 + 2c(5,1))\delta_2(0) = 0$$

$$c(2,2) = \frac{3}{20} - c(2,1) = \frac{3}{20}$$

$$c(3,2) = (\quad 0 \quad + 2c(2,2))/(-3) = \frac{-1}{10}$$

$$c(4,2) = (1c(2,2) + \quad 0 \quad)/1 \quad = \frac{3}{20}$$

$$c(5,2) = (0c(3,2) + 4c(4,2))/(-1) = \frac{-3}{5}$$

$$c(6,2) = (\quad 0 \quad + 2c(5,2))/(-3) = \frac{2}{5}$$

$$c(3,3) = \frac{1}{4} - c(3,1) - c(3,2) = \frac{7}{20}$$

$$c(4,3) = (\quad 0 \quad + \quad 0 \quad)/(4) \quad = 0$$

$$c(5,3) = (0c(3,3) + 4c(4,3))/(2) \quad = 0$$

$$c(6,3) = (\quad 0 \quad + 2c(5,3))\delta_2(0) = 0$$

$$c(4,4) = \frac{3}{10} - c(4,1) - c(4,2) - c(4,3) = \frac{3}{20}$$

$$c(5,4) = (\quad 0 \quad + 4c(4,4))/(-2) = \frac{-3}{10}$$

$$c(6,4) = (\quad 0 \quad + 2c(5,4))/(-4) = \frac{3}{20}$$

$$c(5,5) = \frac{1}{5} - c(5,1) - c(5,2) - c(5,3) - c(5,4) = \frac{11}{10}$$

$$c(6,5) = (\quad 0 \quad + 2c(5,5))/(-2) = -\frac{11}{10}$$

$$c(6,6) = -\sum_{u=1}^{5} c(6,u) = \frac{11}{20}$$

Therefore the solution is

$$\begin{bmatrix} P_{20} \\ P_{11} \\ P_{10} \\ P_{02} \\ P_{01} \\ P_{00} \end{bmatrix} \equiv \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_5 \end{bmatrix} = \frac{1}{20} \begin{bmatrix} 2 & & & & & \\ 0 & 3 & & & & \\ 0 & -2 & 7 & & & \\ 0 & 3 & 0 & 3 & & \\ 0 & -12 & 0 & -6 & 22 & \\ 0 & 8 & 0 & 3 & -22 & 11 \end{bmatrix} \begin{bmatrix} 1 \\ e^{-3t} \\ 1 \\ e^{-4t} \\ e^{-2t} \\ 1 \end{bmatrix} .$$

The same comments which were made in the example of section 4 regarding the ease of finding other derived quantities from the above solution apply here. For example, the distribution function of time to extinction obtained from the above solution is $G(T) = 0$ for $T < 0$ and for $T \geq 0$,

$$G(T) = \sum_{r=0}^{2} P_{r0}(T) = 1 + \frac{1}{20}(6e^{-3T} + 3e^{-4T} - 22e^{-2T}) .$$

REFERENCES

BAILEY, N. T. J. (1957). The Mathematical Theory of Epidemics London: Charles Griffin and Co. Ltd.

BAILEY, N. T. J. (1963). The simple stochastic epidemic: a complete solution in terms of known functions. Biometrika, **50**, 235-240.

GANI, J. (1965). On a partial differential equation of epidemic theory. I. Biometrika, **52**, 617-622.

SISKIND, V. (1965). A solution of the general stochastic epidemic. Biometrika, **52**, 613-616.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) State University of New York at Buffalo Buffalo, New York	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

SOME TOPICS RELATED TO TRANSFORMATIONS, DISTRIBUTION FUNCTIONS AND STOCHASTIC PROCESSES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific. Final.

5. AUTHOR(S) (Last name, first name, initial)

SEVERO, NORMAN C.
SCHILLO, PAUL J.
RODINE, ROBERT H.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
November 1966	83	16

8a. CONTRACT OR GRANT NO. AF33(657)-9885 b. PROJECT NO. 7071 c. 61445014 d. 681304	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) ARL 66-0231

10. AVAILABILITY/LIMITATION NOTICES

1. Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Aerospace Research Laboratories (ARM) Wright-Patterson Air Force Base, Ohio

13. ABSTRACT

This report is divided into four independent sections. Section I contains a theorem giving sufficient conditions for the asymptotic distribution of a standardized location-scale random variable to have the distribution of a power of the random variable, and examples showing that the conditions are not necessary. Section II gives illustrations of the problem of deciding whether or not one random variable is a transform of another; and, in each case in which a transformation of the one random variable into the other is assured, the set of all such transformations is investigated. Section III consists of an example of the notion of robustness of a test as well as a tentative general definition of the concept of robustness of a test, and a brief study of the Kolmogorov metric on the space of location-parameter distributions. Section IV presents simple iterative solutions of special systems of differential-difference equations, in which the constant coefficient matrices are triangular and satisfy conditions sufficient to insure that the solutions involve only exponential terms or terms that are products of linear factors and exponential factors. These methods are applied to the simple stochastic epidemic and to the general stochastic epidemic.

DD FORM 1 JAN 64 1473

Unclassified

Security Classification

Security Classification

14 KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
DISTRIBUTION FUNCTIONS						
TRANSFORMATIONS						
ROBUSTNESS						
EPIDEMIC THEORY						
LOCATION-SCALE DISTRIBUTIONS						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ______________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ______________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ______________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.

Security Classification