Statistical Science
2002, Vol. 17, No. 2, 209–219

On Exchangeable, Causal and Cascading Failures

Dennis V. Lindley and Nozer D. Singpurwalla

Abstract. This paper is addressed to engineers and statisticians working on topics in reliability and survival analysis. It is also addressed to designers of network systems. The material here is prompted by problems of infrastructure assurance and protection. Infrastructure systems, like the internet and the power grid, comprise a web of interconnected components experiencing interacting (or dependent) failures. Such systems are prone to a paralyzing collapse caused by a succession of rapid failures; this phenomenon is referred to as "cascading failures." Assessing the reliability of an infrastructure system is a key step in its design. The purpose of this paper is to articulate on aspects of infrastructure reliability, in particular the notions of chance, interaction, cause and cascading.

Following a commentary on how the term "reliability" is sometimes interpreted, the paper begins by making the argument that exchangeability is a meaningful setting for discussing interaction. We start by considering individual components and describe what it means to say that they are exchangeable. We then show how exchangeability leads us to distinguish between chance and probability. We then look at how entire networks can be exchangeable and how components within a network can be dependent. The above material, though expository, serves the useful purpose of enabling us to introduce and make precise the notions of causal and cascading failures. Classifying dependent failures as being either causal or cascading and characterizing these notions is a contribution of this paper. The others are a focus on networks and their setting in the context of exchangeability.

A simple model for cascading failures closes the paper. A virtue of this model is that it enables us to make the important claim that causal failures are more deleterious to infrastructure reliability than cascading failures. This claim, being contrary to a commonly held perception of network designers and operators, is perhaps the key contribution of this paper.

Key words and phrases: Chance, dependence, infrastructure, interaction, network, probabilistic causality, reliability, survival.

1. PREAMBLE

The scenario of infrastructure protection has created an opportunity for interaction between engineers, network designers and statisticians on a matter of recent concern. However, it has also created some problems of communication, interpretation and focus. One such problem pertains to an appreciation of the nature of "reliability." Whereas reliability has been clearly defined in the statistical literature (to include engineering outlets such as the *IEEE Transactions on Reliability*) it remains conversational and nebulous to many of the designers, builders and operators of infrastructure systems. This has been the experience of one of

Nozer D. Singpurwalla is a Professor at George Washington University, Washington, DC 20052 (e-mail: nozer@research.circ.gwu.edu). Dennis V. Lindley is Professor of Statistics, "Woodstock," Quay Lane, Minehead, TA24 5QU, U.K.

us. Questions are often raised, such as: what do we mean by the word "reliability"? Why should reliability have a quantitative import? Is reliability not a chance, and if so, what is the difference between chance and probability?

To some, reliability is simply an assertion: "if an item works, then it is reliable; otherwise, it is unreliable." To others, reliability cannot have a universal definition, because it should be context dependent. For example, whereas the reliability of an automobile represents its ability to provide transportation on demand, the reliability of a power system must reflect its ability to deliver power at specified thresholds for specified periods of time. The view that reliability is simply an assertion about the functioning or not of an item is naïve. For one thing, it does not take into account the random nature of failures. Indeed, it is a recognition of the randomness of failure which makes it possible for us to quantify reliability as a probability, and this in turn enables us to come to terms with a context-free definition of reliability. More important, once the stochastic nature of failure is acknowledged, notions such as "chance," "interaction" and "cascading," terminology that is the mainstay of infrastructure assurance can be made precise.

This paper is chiefly addressed to engineers and statisticians working on problems of reliability and survival analysis. However, with a focus on infrastructure as a point of discussion, the paper may also be of interest to network theorists. They may find our view of looking at networks within the framework of exchangeability valuable. Furthermore, they may find our conclusions about the deleterious nature of causal and cascading network failures intriguing. Sections 2 and 4 of the paper could also appeal to statisticians interested in foundational issues, independent of specific applications.

1.1 Chance and Probability

To most engineers, and for that matter many mathematicians and statisticians, the notions of chance and probability are synonymous. However, this need not be universally so. Under the setting of exchangeability, described below, chance is an objective feature independent of an individual (or a group of individuals acting as one). The objective feature is a property of something that is akin to what is known as a "collective" [cf. Von Mises (1957)]. Collectives are conceptual entities that involve a kind of homogeneity, and sequences of infinite size. By contrast, "probability" connotes a flavor of individuality, in the sense that it reflects an individual's disposition to bet. Thus probability is specific to an individual, or a group of individuals acting as one; that is, probability is subjective. This is the view of probability that we shall adopt. However, not all share this view, so that for them chance and probability may indeed be synonymous. The view of probability as subjective does pose a difficulty in public discourse, and this could be a reason for its lack of universal appeal.

The distinction between chance and probability is made transparent by a theorem of de Finetti [cf. Lindley and Phillips (1976)] on an infinite number of unknown quantities that are "exchangeable"; see Section 2. Exchangeability, as a judgment made by an individual or a group of individuals, plays a central role in reliability and survival analysis. Its practical import is that it enables one to make predictions based on observations of observables that are judged exchangeable with the unobservables, for example, life testing. Clearly, relating the observed to the unobserved must connote an underlying notion of dependence. Indeed, the judgment of exchangeability is, de facto, a statement of dependence.

To summarize, the quantification of reliability by probability enables us to articulate the notions of chance, dependence, exchangeability and independence. Furthermore, as we shall see later, the notion of dependence gives birth to that of causality, and this in turn helps us characterize cascading. The quantification of reliability therefore provides a basic requirement of science, namely, an ability to measure.

2. EXCHANGEABLE COMPONENTS

A network consists of a number of components operating together according to a specified architecture; series and parallel systems are simple examples. However, before considering their joint action we look at a single component in the network. This component will ultimately fail, and on its installation you will have beliefs, expressed through *probability*, about its time to failure. Denote by $F(t)$ your probability that it will still be functioning at time t, so that if X is its lifetime, $F(t) = \mathcal{P}(X > t)$, and $-dF(t)/dt = f(t)$ is the *probability density* of X. Our notation is contrary to convention, in that $F(t)$ generally denotes $\mathcal{P}(T \leq t)$. Here, $\mathcal{P}$ denotes probability. Two points about $F(t)$ need to be emphasized; first, that as a probability, it expresses your opinion, rather than the actual performance, in any sense, of the component; second, that it depends not only on the component but also on the conditions under which it is used. $F(t)$ is known as the *survival function* of the component for a *mission time* t, $t \geq 0$. For reasons that will become clear in the sequel, we do not refer to $F(t)$ as the "reliability" of the component; that is, we will be making a distinction between reliability function and the survival function. $F(t)$ may be interpreted as a measure of the desirability of the component as viewed by an individual, namely us. The universality of this interpretation comes from the fact that X need not represent lifetime; X could represent, for example, the amount of power delivered, and $\mathcal{P}(X > t)$ represents our uncertainty about the delivered power being greater than a threshold t.

It is usual, in the context of reliability theory, to think of the component as coming from a batch of similar components, each of which might have replaced it in the network with similar results. It is necessary to express the notion of similarity in a mathematical form and a way to do this is through the concept of *exchangeability*, mentioned briefly in Section 1.1. For n components from the batch, let $f(t_1, t_2, \dots, t_n)$ be your probability density that they will fail at times $t_1, t_2, \dots, t_n$; then you are said to judge them as exchangeable if this probability is invariant under any permutation of the subscripts. For example, with two components, your probability that the first will fail at t_1 and the second at t_2 is the same as that for the first at t_2 and the second at t_1; that is, $f(t_1, t_2) = f(t_2, t_1)$. Since exchangeability is defined in terms of probability, and since to us here probability, as an opinion, is a judgment, exchangeability is therefore also a judgment; see Section 1.1. A famous theorem of de Finetti (1938) says that for an infinite batch, or, more practically, a very large batch, exchangeability implies that your beliefs have a structure in which the lifetimes are believed to be independent and identically distributed (i.i.d.) with a distribution, that is unknown to you but about which you have beliefs, reflected through probability. Furthermore, this distribution is the limit of your beliefs about the empirical distribution of the observed lifetimes; see, for example, Bernardo and Smith (1994), page 177. De Finetti's theorem involves some rather complicated mathematics and it is usual and useful to think about it in terms of a parameter θ that indexes all distributions on the real line, where we can write, taking some liberties with precision,

$$f(t_1, t_2, \dots, t_n) = \int \prod_i f(t_i|\theta) f(\theta)\, d\theta \tag{1}$$

with density $f(t_i|\theta)$ of individual lifetimes, the same for all, the product expressing their independence, given θ, and a density $f(\theta)$ for the parameter θ; θ could be a scalar or a vector. Exchangeability therefore produces the usual assumption of i.i.d. lifetimes involving a parameter θ and a "prior" distribution for θ. Under reasonable conditions, as n increases, the parameter concentrates around a value, θ_0 say, and $f(t|\theta)$ tends to a limit $f(t|\theta_0)$ which is the empirical density; θ_0 is often referred to as the true value of θ. This limit is called a *chance*, and the *reliability* of the component for a mission of time τ, is $\int_\tau^\infty f(t|\theta_0)\, dt$. In writing the above we have made a distinction between reliability and the survival function through chance and probability, respectively. Thus, in the context of exchangeability and de Finetti's theorem, we make the claim that *reliability is a chance*, not a probability.

While θ indexes all distributions on the real line, it is usual to use a small subclass of such distributions; for example, exponential distributions with mean $1/\theta$, and we speak of the subclass as providing a *failure model* for the lifetimes. This greatly simplifies the mathematics but causes trouble if the empirical distribution or chance departs seriously from the assumed exponential form. In making a simplification of this type we have to recognize that the judgment (here exponential) of $f(t|\theta)$ is a judgment by you, not an objective statement. The objective quantity is the limit of the empirical distribution function, which in principle can never be observed. Even this is only objective in the sense that it is shared by all who make the exchangeability judgment. Diaconis and Freedman (1980) have trouble with the synthesis about objective quantities; see Section 6, Remark 1 of their paper. The individuality of $f(\theta)$, the "prior," however, is recognized and is often given as a reason for rejecting the position taken here. The choice of suitable failure models occupies much literature in the statistical theory of reliability.

2.1 Dependence and Independence

Consider now the case of two components from a batch which are judged exchangeable; the ideas about to be presented extend to any number. It is important to distinguish two types of probability statements about their lifetimes t_1, t_2. First, there is your joint density $f(t_1, t_2)$ expressing your belief that the first will fail at t_1, the second will fail at t_2; second, there is your joint density, given θ, $f(t_1|\theta) f(t_2|\theta)$. In the latter, t_1 and t_2 are judged by you to be *independent* (and identically distributed), were you to know θ. In the former, they are judged to be *dependent* because of the linkage provided by θ, which is unknown and expressed in (1). This linkage allows learning in the sense that your judgment about t_2 will be affected by observing the value of t_1. In principle, with θ unknown, observing t_1 enables you to revise your judgments about θ, and this in turn enables you to revise your judgments about t_2. Were θ to be assumed known, no such learning occurs because of the independence. Thus the notions of dependent and independent lifetimes are statements about learning. Whenever there is learning our probabilities change; otherwise they do not. Again the exponential distribution provides an illustration. Here, suppose that for $i = 1, 2$,

$$f(t_i|\theta) = \theta \exp(-\theta t_i),$$

and thus by (1),

$$f(t_1, t_2) = \int \theta^2 \exp[-\theta(t_1 + t_2)] f(\theta)\, d\theta.$$

For the convenient, simple case, with $f(\theta) = u \exp(-\theta u)$, easy calculation gives

$$f(t_1, t_2) = 2u/(t_1 + t_2 + u)^3$$

and

$$f(t_2|t_1) = 2(t_1 + u)^2/(t_1 + t_2 + u)^3,$$

demonstrating how your belief about t_2 is affected by observation of t_1. Note that $f(t_2|t_1)$ is the conditional density of t_2 were t_1 to be observed; $f(t_2|t_1) = f(t_1, t_2)/f(t_1)$. In particular, notice that this density is not exponential and in the upper tail behaves like t_2^{-3} rather than like $\exp(-\theta t_2)$.

While it is usual to assume some small class of distributions, like the exponential or the Weibull, difficulties can arise if the true distribution, that is, the empirical limit, were not a member of the assumed class. If the chance distribution is truly exponential, then your opinions about θ, expressed through $f(\theta|t_1, t_2, \ldots, t_n)$ will concentrate around the true value. If the exponential is assumed but truth lies elsewhere, then, intuitively speaking, the situation seems to be that θ will tend to the value that makes the true distribution as close to the exponential with that value, in the sense of Kullback and Leibler (1951) distance, as possible. One strategy is to make the model as large as computational facilities will allow, keeping in mind considerations of parsimony (i.e., Occam's razor). Another factor to be borne in mind is that in practice one rarely observes a large number of similar components, so that the empirical distribution is poorly determined. This probably accounts for successes achieved with models as restricted as the exponential. Although it is true that repetitions are rarely sufficiently large to permit a reasonable assessment of the chance distribution, those with considerable experience with many networks and their components can obtain a fairly good idea that the chance distribution belongs to a class of densities, like the Weibull. It is this sort of collective experience that encourages the use of standard distributions.

3. EXCHANGEABLE NETWORKS

We now pass from a single component, taken from a batch, to a network of, say, two components, each component not necessarily from the same batch. They will have lifetimes t_1 and t_2 dependent on the network architecture and the external conditions, and you will have a density $f(t_1, t_2)$ for them, expressing your beliefs about when they might fail. You may judge them exchangeable, in which case t_1 and t_2 will be dependent, but conditionally independent given some θ, and the considerations of Section 2 will apply. However, the general case need not restrict itself to intercomponent exchangeability. It would be more reasonable to suppose that a collection of similar networks, being taken from a batch of networks, is judged exchangeable. If so, there is a bivariate form of de Finetti's theorem using $f(t_1, t_2|\theta)$ and $f(\theta)$. For example, with two exchangeable networks with failure times t_{ij} for component i in network j, $i, j = 1, 2$,

$$f\big((t_{11}, t_{21}), (t_{12}, t_{22})\big) = \int f(t_{11}, t_{21}|\theta) f(t_{12}, t_{22}|\theta) f(\theta)\, d\theta.$$

In general, with n exchangeable two-component networks, the bivariate analogue of (1) is

$$f\big((t_{11}, t_{21}), \ldots, (t_{1n}, t_{2n})\big) = \int \prod_i f(t_{1i}, t_{2i}|\theta) f(\theta)\, d\theta, \tag{2}$$

reflecting independence of the vectors (t_{1i}, t_{2i}), given θ, $i = 1, \ldots, n$, and $f(t_{1i}, t_{2i}|\theta)$ being the same for all i.

Under reasonable conditions, as n increases, the parameter θ concentrates around a value, say θ_0, and $f(t_1, t_2|\theta_0)$ is the limit of the bivariate empirical density of (t_{1i}, t_{2i}), $i = 1, 2, \ldots$; this limit is the chance. The subclass of distributions used to represent $f(t_1, t_2|\theta_0)$ is then a *bivariate* failure model for the lifetimes.

From $f(t_1, t_2)$, your belief about the failure time of the network can be calculated, and from $f(t_1, t_2|\theta)$ the same belief, given θ, can be found. These will depend on the network's architecture; for example, if the two components are in series (parallel), then the time to network failure is the smaller (larger) of t_1, t_2. For the case $n = 2$ of (2), while given θ, the network lifetimes are judged independent, it does not imply, that with θ known, the component lifetimes within each network are independent. That is, $f(t_1, t_2|\theta)$ may not factor as $f(t_1|\theta)$ and $f(t_2|\theta)$. For example, $f(t_1, t_2|\theta)$ could be any one of the several bivariate distributions in the literature. Thus, it is entirely possible, that when the lifetimes of two networks are judged exchangeable, these lifetimes can be independent given θ, even though the lifetimes of each component within each network are dependent. With θ unknown, the lifetimes of the network will of course be dependent, assuming that they are exchangeable. It is rarely satisfactory to say that two quantities like t_1 and t_2 are independent. It is also necessary to state the conditions under which your probability for t_1 and t_2 is being discussed. Thus there is a real difference between their being independent, given θ, and being independent given only the background knowledge of the network and its components.

There are several ways in which, even with θ known, dependence of component lifetimes within a network can arise. We describe here two possibilities. However, we must first emphasize the fact that for networks with many components, even as few as five, such component dependencies can create havoc vis-a-vis the determination of the network's reliability function; indeed, for such networks the "reliability polynomial," which is defined for networks with independent components (given θ), does not exist! The first of the above mentioned two possibilities is that circumstances outside the network may put a strain on it; this strain could be shared by both components rendering their lifetimes correlated. For example, with the power line network, a storm (common shock) could bring down several trees in different parts of the region, causing several lines to break. For such scenarios, the bivariate exponential distribution of Marshall and Olkin (1967) wherein $\underline{\theta} = (\lambda_1, \lambda_2, \lambda_{12})$,

$$\mathcal{P}(X_1 \geq t_1,\ X_2 \geq t_2|\underline{\theta}) = \exp[-\lambda_1 t_1 - \lambda_2 t_2 - \lambda_{12}\max(t_1, t_2)] \tag{3}$$

is an attractive model for approximating the chance density $f(t_1, t_2|\theta)$. Marshall and Olkin introduced a model for a system with two components in which there are three types of external shock to the system. Type 1(2) affects component 1(2) only, whereas type 3 affects both components simultaneously. The shocks are generated by three independent Poisson processes having rates λ_1, λ_2 and λ_{12}, respectively. The lifetimes X_1 and X_2 of the two components are dependent, even with $\underline{\theta}$ known, since the marginals $\mathcal{P}(X_i \geq t_i) = \exp(-(\lambda_i + \lambda_{12})t_i)$, $i = 1, 2$, obtained by letting t_j, $j \neq i$, in (3) become zero, cannot be multiplied to yield (3). Why this dependence? It stems from the fact that upon observing t_1 (or t_2), our assessment of X_2 (or X_1) changes on the grounds that learning about a failure at t_1 (or t_2) leads us to incorporate the belief that the common shock could have been the cause of failure. Knowing $\underline{\theta}$ does not tell us when the common shocks occur, but knowing t_1 (or t_2) causes us to believe that the common shock could have occurred there. A good summarization of this bivariate exponential distribution is in Barlow and Proschan (1975), page 128.

For the second possibility leading to intercomponent lifetime dependency within a network, we consider the case of two components in parallel. If one component fails, the network still functions, but the surviving component is subjected to an increased stress, and thus the lifetimes have a positive association. Paired organs is a good example from the biological sciences, and fiber bundles is an example from the physical sciences. As a simple scenario, suppose that, given θ, the lifetimes of the individual components are independent with $\theta\exp(-\theta t)$ as a common density; that is, the failure model is an exponential. Now suppose that when one of the two components fails, there is an increased stress on the surviving component so that θ doubles to 2θ and its remaining lifetime is exponential with density $2\theta\exp(-2\theta t)$; see Figure 1. Simple calculations show that the joint density of the lifetimes at t_1 and t_2 is $2\theta^2\exp(-2\theta t_2)$ for $t_1 < t_2$, and $2\theta^2\exp(-2\theta t_1)$ for $t_2 < t_1$; t_i, $i = 1, 2$, is the lifetime of the ith component. The time to failure of the network has density at u, of the form $4u\theta^2\exp(-2\theta u)$; its mean time to failure is $1/\theta$. The model for dependent lifetimes given above is a restricted case of the bivariate exponential distribution of Freund (1961). Here dependency is because upon the failure of the first component, we learn of the time at which that surviving component experiences an increased stress, and this in turn changes our assessment of the survivability of the second component. More details about this can be found in Freund (1961), who also develops expressions for the marginal densities and the joint moments.

Examples of other bivariate distributions used as chance distributions in survival analysis are summarized by Oakes (2001) and in engineering reliability

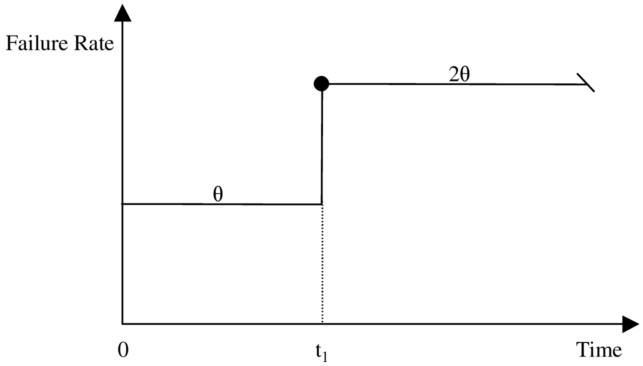

FIG. 1. *Failure rate of the second component to fail, in Freund's model.*

by Martz and Waller (1982) and by Henley and Kumamoto (1981), who have other models of common cause failures.

4. CAUSAL, CASCADING AND INTERACTING FAILURES

The notion of dependence as a probability changing phenomenon (due to learning) has been well incorporated and developed in reliability theory and in survival analysis. Independence, always conditional (given some parameter) simplifies the assessment of probability, and is often an idealization. Several stochastic failure models incorporating dependence have been introduced; the ones by Freund (1961) and by Marshall and Olkin (1967) are two examples. However, in practice, we also encounter terms such as "the cause of failure," "interaction" and "cascading failures." These terms do not appear to have been well articulated within a mathematical framework. In this section we outline some preliminary thoughts that may help alleviate some of these shortcomings. We start with causality, trace its philosophical roots, and then develop some mathematical relationships. We then argue that the models by Marshall and Olkin and by Freund are indeed models of causal failure. The notion of causality with an added caveat leads us to the notion of cascading. Consequently, a modified version of the model by Freund paves the way towards a probabilistic characterization of cascading failures.

4.1 Causality and Models of Causal Failure

4.1.1 *Philosophical background.* The notion of *causality*, or *causation* has been elusive, from the days of Hobbes, Hume and Newton to the modern day philosophers such as Suppes and Salmon; see, for example, the excellent overview by Holland (1986). Four principal conceptions of causation have been proposed to address the question "What is a causal relationship?" These are the materialist (or dynamical), the spiritualist, the rationalist (or apriorist), and the phenomenalist (or positivist); see Krajewski (1982), pages 223–235.

The *materialist* notion was enunciated by Hobbes and was adopted by Newton in his *Principia.* Here the causal relation is the connection between two bodies $\mathcal{A}$ and $\mathcal{B}$, in which $\mathcal{A}$ acts on $\mathcal{B}$ and makes changes to it. The action is connected with the transmission of motion; to Newton, cause was a force. This view underwent a change via physicists like Helmholtz, Ostwald, Planck and Einstein, so that causal relations were identified with the transmission of energy, and cause is the principle of conservation of energy. The contemporary view takes into account nonphysical phenomena (biological and sociological) so that causation is considered as an action involving a transfer of energy or of information.

The *spiritualist* notion traces its roots to the Bishop of Berkeley and to Aquinas. Here cause is identified as a spiritual being that by virtue of its will causes a change in another being (spiritual or material). Thus, it is only God that "is the true cause." The *rationalist* notion views a causal relation as a special case of the logical relation between reason and consequence. Thus Descartes considers cause to be the reason for the existence of a thing. This view was also held by Spinoza and by Leibniz.

By far the most common and the most discussed conception is the phenomenalist one. It is due to Hume who sees a causal relation as one that can only be drawn from experience. To Hume (1748), cause is

> ... an object precedent and contiguous to another, and so united with it, that the idea of one determines the mind to form the idea of the other.

That is, the frequent observation that $\mathcal{A}$ is followed by $\mathcal{B}$ associates in our mind the notion that when $\mathcal{A}$ appears we expect $\mathcal{B}$ also to appear. Thus, the connection between events is of a psychological, associative nature. Hume's psychologism has been the topic of much discussion. However, his basic notion of succession of events has been accepted, and it has been enriched by others, namely, John Stuart Mill (1843), who claimed that a causal relationship between $\mathcal{A}$ and $\mathcal{B}$ can be ascertained, not only if $\mathcal{A}$ is always followed by $\mathcal{B}$, but when, in addition, we know that this will continue in the future in all situations, regardless of other circumstances. That is, $\mathcal{B}$ must follow $\mathcal{A}$ unconditionally. This unconditionality must be drawn from sources other than experience, for which Mill appeals to science as a repository of mankind's extensive experience. Further refinements to Hume's thesis of succession (temporal), and Mill's thesis of the perpetual are that if events are separated in space, one needs to look at intermediate events, events that may not be perceived by direct observation, but inferred by knowledge of physics and psychology; that is, a *spatio-temporal contiguity*.

Of these four conceptions, it is the one by Hume, that is, the phenomenalist conception, that has motivated a quantification of causality, one deterministic and the other probabilistic. The former, which is overviewed in the Appendix, is subsumed by the latter under an additional condition. Probabilistic causality enables us to identify failure modes that are causal, and its modification enables us to characterize failures that cascade. The material in the Appendix (which if included in the text would be distracting) enables us to put the material of Section 4.1.2 in its proper context.

4.1.2 *Probabilistic causality.* A probabilistic approach to causality has been developed by Reichenbach, Good, and Suppes; see, for example, Salmon (1980), pages 50–74. We outline below the version of Suppes (1970). In the context of time series this version is known as Wiener causality [cf. Granger (1969)]. Here, $\mathcal{D}$ is a *prima facie probabilistic cause* of $\mathcal{E}$, denoted $\mathcal{D} \overset{P}{\rightarrow} \mathcal{E}$, if:

1. $\mathcal{D}$ occurs before $\mathcal{E}$ (in time).
2. $\mathcal{P}(\mathcal{D}) > 0$.
3. $\mathcal{P}(\mathcal{E}|\mathcal{D}) > \mathcal{P}(\mathcal{E})$.

Condition (3) implies that a cause is a probability raising event. Contrast this with the notion of dependence wherein all that matters is a change in probability. We note that (3) is also a consequence of deterministic causality (see the Appendix), since there $\mathcal{D} \overset{C}{\rightarrow} \mathcal{E} \Rightarrow \mathcal{P}(\mathcal{E}|\mathcal{D}) > \mathcal{P}(\mathcal{E})$. Thus we have the theorem.

THEOREM 1. *If* $\mathcal{D}$ *occurs before* $\mathcal{E}$, *and if* $0 < \mathcal{P}(\mathcal{D}) < 1$, *then*

$$\mathcal{D} \overset{C}{\rightarrow} \mathcal{E} \quad \Longrightarrow \quad \mathcal{D} \overset{P}{\rightarrow} \mathcal{E}.$$

The converse is not true. Thus the notion of probabilistic causality is a weakening of the notion of deterministic causality, and as such is a broadening of this class.

From condition (3) it is easy to see that if $\mathcal{P}(\mathcal{E}) > 0$, then $\mathcal{D} \overset{P}{\rightarrow} \mathcal{E} \leftrightarrow \mathcal{P}(\mathcal{D}|\mathcal{E}) > \mathcal{P}(\mathcal{D})$, and that under probabilistic causality, $\mathcal{D}$ cannot be disjoint from $\mathcal{E}$. However, $\mathcal{D}$ can be completely contained in $\mathcal{E}$, or $\mathcal{D}$ and $\mathcal{E}$ can intersect, so that the first and the middle illustrations of Figure 6 correspond to the case $\mathcal{D} \overset{P}{\rightarrow} \mathcal{E}$, but not the case $\mathcal{D} \overset{C}{\rightarrow} \mathcal{E}$.

Suppes' three conditions define the cause to be *prima facie*, because the cause is only an apparent cause. Suppes declares a cause to be a *genuine* cause, if it is a prima facie cause that cannot be shown as being a "spurious cause." A (prima facie) cause $\mathcal{D}$ is said to be a *spurious cause* of $\mathcal{E}$ if and only if there exists a cause $\mathcal{S}$ where:

(i) $\mathcal{S}$ occurs before $\mathcal{D}$.
(ii) $\mathcal{P}(\mathcal{D}, \mathcal{S}) > 0$.
(iii) $\mathcal{D}$ is a prima facie cause of $\mathcal{E}$.
(iv) $\mathcal{P}(\mathcal{E}|\mathcal{D}, \mathcal{S}) = \mathcal{P}(\mathcal{E}|\mathcal{S})$.
(v) $\mathcal{P}(\mathcal{E}|\mathcal{D}, \mathcal{S}) \geq \mathcal{P}(\mathcal{E}|\mathcal{D})$.

Thus a spurious cause is a prima facie cause that can be explained away by conditioning on an earlier event (or a common cause) that accounts as well for the conditional probability of the effect. Clearly, since there could be an inexhaustible number of $\mathcal{S}$'s that could render a prima facie cause spurious, one can rarely ascertain that a particular cause is genuine. It is because of arguments such as this, that Suppes's notion of probabilistic causality has been, to some philosophers, unsatisfactory; see Hesslow (1976, 1981).

However, defenders of this theory, such as Rosen (1978) maintain that an assertion of a causal relationship depends on what information is available and upon a conceptual framework relative to which the causal relations are postulated. It is possible that the conceptual framework is inadequate or partial, and thus there is always the possibility of discovering a better explanation of causal relationships. Thus Suppes' probabilistic theory presupposes the view that there are no ultimate causes and that to assume otherwise is sheer dogmatism. Consequently, all the probability statements in this section must be rewritten, to include the background information (or the conceptual framework $\mathcal{H}$); for example, (2) should be written as $\mathcal{P}(\mathcal{D}; \mathcal{H}) > 0$ and (v) as $\mathcal{P}(\mathcal{E}|\mathcal{D}, \mathcal{S}; \mathcal{H}) \geq \mathcal{P}(\mathcal{E}|\mathcal{D}; \mathcal{H})$.

The bivariate exponential distribution of Marshall and Olkin discussed in Section 3 provides an illustration of probabilistic causality. Let $\mathcal{E}$ be the event that $X_2 = t_2$, and $\mathcal{D}$ the event that $X_1 = t_1$, where $t_1 < t_2$. Then it is easy to verify that for $t_1 > \frac{1}{\lambda_{12}} \ln \frac{\lambda_1+\lambda_{12}}{\lambda_1}$, the three conditions of prima facie causality are satisfied so that $\mathcal{D}$ is a prima facie probabilistic cause of $\mathcal{E}$. However, $\mathcal{D}$ is not a genuine cause. If $\mathcal{S}$ denotes the event that the common shock of this distribution occurs at t_1, then $\mathcal{E}$ is independent of $\mathcal{D}$, given $\mathcal{S}$. Indeed, in this particular case, $\mathcal{S}$ is a genuine probabilistic cause of $\mathcal{E}$. Thus we claim that Marshall and Olkin's bivariate distribution is a model of causal failure for $\mathcal{A}$ with $T_2 = t_2$ as the event, $T_1 = t_1$ as a prima facie cause and $\mathcal{S}$ as a genuine cause, for all values of t_1 greater than a specified constant.

4.1.3 *Causes of effects and effects of causes.* The deterministic and probabilistic thinking outlined above is concerned with assessing the *cause of an effect.* Such assessments are useful in medical diagnosis and in maintenance management. In the latter context, the effect could be the network's failure, and a prima facie cause the failure of one or more of its nodes. The comment by Cox (1986) on Holland (1986) raises some noteworthy philosophical and statistical issues. By contrast, when we are addressing issues pertaining to the efficacy of a treatment or an engineering design, interest centers around assessing the *effects of a cause*. For example, we may want to know what could happen to a network if one or more of its nodes were to be disabled, or if several nodes were collapsed into one.

Bayes' law is a vehicle for assessing the causes of effects. Thus for example, if $\mathcal{C}$ or $\mathcal{D}$, or both $\mathcal{C}$ and $\mathcal{D}$, are probabilistic causes of an event $\mathcal{E}$, then $\mathcal{P}(\mathcal{C}|\mathcal{E}) \propto \mathcal{P}(\mathcal{E}|\mathcal{C})\mathcal{P}(\mathcal{C})$; similarly $\mathcal{P}(\mathcal{D}|\mathcal{E})$. On the other hand, assessing the effects of causes has proved to be a difficult endeavor, especially in the medical context, since every subject under study can experience a treatment (i.e., a cause) or not. Much has been written on this topic in the statistical literature. This we have barely touched upon since our focus here is on probabilistic modeling. A comprehensive synthesis of a broad spectrum of issues involving statistical techniques is by Cox (1992).

4.1.4 *Other approaches to causality.* To some, such as Pearl (2000), the approaches to causality discussed here seem inadequate. Perhaps his ideas are easily the best we have because no definition of causality can be expected to embrace all the meanings that we associate with the word. Like most words in the English language, the term "cause" is imprecise, whereas any definition of the term is precise. The same can also be said of the word "probable"; however, the notion of probability has been made precise in one of several ways.

4.2 Cascading and Models of Cascading Failures

What are cascading failures and how are they different from causal failures? Before addressing this question, it is helpful to overview the distinction between dependent and causal failures. Whereas dependence is characterized in terms of probability changing events, causality is characterized, among other things, in terms of probability increasing events. Furthermore, with dependence, there being no consideration of a time sequencing of occurrences, there is an interchangeability of events. Thus if $\mathcal{D}$ is dependent on $\mathcal{E}$, then $\mathcal{E}$ is dependent on $\mathcal{D}$. With causality there is a time ordering so that if $\mathcal{D}$ is the cause of $\mathcal{E}$, then $\mathcal{E}$ cannot be the cause of $\mathcal{D}$. Engineers refer to this time ordering as *dependence with dynamics*.

To distinguish causal failures from cascading failures, the first thing to note is that in a causal failure model, simultaneous failures are possible. For example, in Marshall and Olkin's model, an occurrence of the common shock (the genuine cause) is followed by the simultaneous failure of both components. Under the notion of cascading that will be introduced later, there is a sequence of failures, one followed by the other, but within a specified time; no simultaneous failures are allowed. The scenario here parallels that of a domino effect. That is, the falling of a domino causes its neighbor to fall, but only if the neighbor is within striking distance of the falling domino. If the dominos are too far apart, a falling domino will not have any effect on its neighbors. Thus with cascading failures, the failure of one component is followed by that of its neighbor, but within a specified time, which we shall call *critical time*. If the failure of a component causes its neighbor to fail, but after the critical time has elapsed, then such failures are causal, not cascading.

Freund's model, discussed at the end of Section 3, provides a suitable framework for developing models of cascading failures. Recall that in this model the failure of the first component permanently changes the parameter θ—which is the failure rate of the surviving component—to 2θ. This increase in the failure rate, shown in Figure 1, increases the probability of failure of the surviving component. Thus, if $X_1 = t_1$ denotes the event that the first component to fail, fails at time t_1, and if $X_2 = t_2$, $t_1 < t_2$, denotes the event that the surviving component fails at time t_2, then the event $(X_1 = t_1)$ is a prima facie cause of event $(X_2 = t_2)$, and thus Freund's model is also a description of causal failures.

Now suppose that we modify Freund's model such that the failure rate of the surviving component changes

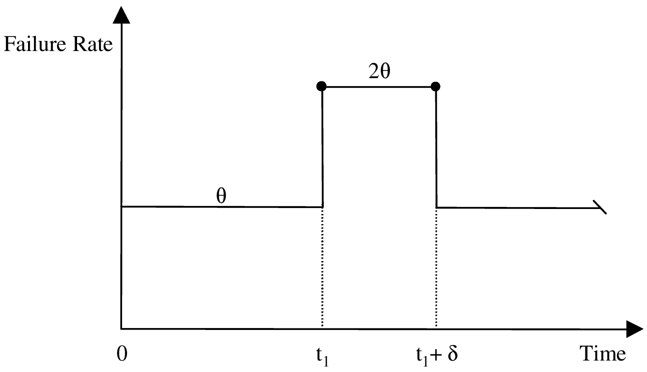

FIG. 2. *Failure rate of the second component to fail in a model for cascading failures.*

at t_1 from θ to 2θ, but at time $t_1 + \delta$ it reverts back to θ, as illustrated in Figure 2. The quantity $\delta > 0$ is the critical time (or the threshold time), and the ensuing model is a description for a cascading failure. The choice of what value to choose for δ is subjective, though a possible strategy is to let δ be the time it takes to restore the failed component to operational status.

For the setup of Figure 2 we can verify [see Swift (2001)] that the joint density at t_1 and t_2, for $t_1 < t_2$ is

$$\begin{array}{ll} 2\theta^2 e^{-2\theta t_2} & \text{if } t_2 < \delta, \\ 2\theta^2 e^{-2\theta t_2} & \text{if } t_1 < t_2 < t_1 + \delta, \\ \theta^2 e^{-\theta(t_1+t_2+\delta)} & \text{if } t_2 > t_1 + \delta. \end{array}$$

Similarly, the joint density for $t_2 < t_1$ is

$$\begin{array}{ll} 2\theta^2 e^{-2\theta t_1} & \text{if } t_1 < \delta, \\ 2\theta^2 e^{-2\theta t_1} & \text{if } t_2 < t_1 < t_2 + \delta, \\ \theta^2 e^{-\theta(t_1+t_2+\delta)} & \text{if } t_1 > t_2 + \delta. \end{array}$$

The time to failure of the system with cascading failures has density at u of the form

$$f_C(u) = \begin{cases} 4\theta^2 u e^{-2\theta u}, & u < \delta, \\ 4\theta^2 \delta e^{-2\theta u} + 2\theta e^{-\theta(u+\delta)} & \\ \quad - 2\theta e^{-2\theta u}, & u \geq \delta. \end{cases}$$

Its survival function at u is of the form

$$S_C(u) = \begin{cases} 2\theta u e^{-2\theta u} + e^{-2\theta u}, & u < \delta, \\ 2\theta\delta e^{-2\theta u} + 2e^{-\theta(u+\delta)} - e^{-2\theta u}, & u \geq \delta; \end{cases}$$

thus its failure rate at u is

$$h_C(u) = \begin{cases} 4\theta^2 \dfrac{u}{2\theta u + 1}, & u < \delta, \\ \dfrac{2\theta(2\theta\delta e^{-\theta u} + e^{-\theta\delta} - e^{-\theta u})}{2\theta\delta e^{-\theta u} + 2e^{-\theta\delta} - e^{-\theta u}}, & u \geq \delta. \end{cases}$$

The mean time to system failure is $\frac{1}{\theta} + \frac{1}{2\theta} e^{-2\theta\delta}$; this is larger than $\frac{1}{\theta}$, which is the mean time to failure under the causal model of Freund. This is to be expected. We also note that as $\delta \uparrow \infty$, the mean time to system failure is $\frac{1}{\theta}$, and that as $\delta \downarrow 0$, the mean is $\frac{3}{2\theta}$, the mean time to failure of a parallel redundant system with two independent, exponentially distributed life-lengths. Thus for parallel redundant systems whose component life-lengths have exponentially distributed life-lengths, cascading failures result in a larger mean time to failure than causal failures, but a lower mean time to failure than that under independent failures.

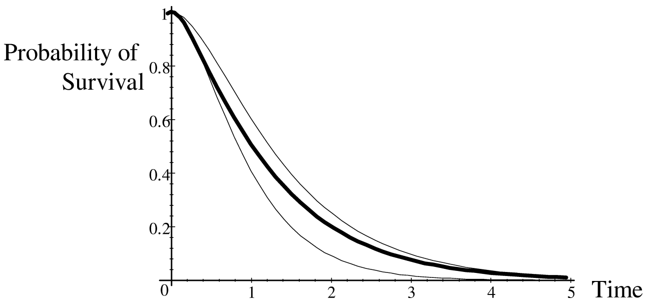

FIG. 3. *A comparison of survival functions under causal* (*lower line*), *cascading* (*thick line*) *and independent* (*upper line*) *failures* (*when $\theta = 1$ and $\delta = 0.25$*).

To see if the above characteristic is also displayed by the survival (or reliability) function, we recall that under Freund's model of causal failures, the density of the system failure at u is of the form $4\theta^2 u e^{-2\theta u}$, so that its survival function, say $S_F(u) = 2\theta u e^{-2\theta u} + e^{-2\theta u}$, and its failure rate is $h_F(u) = 4\theta^2 u/(2\theta u + 1)$. Similarly, for a two-component parallel redundant system with independent exponentially distributed life-lengths, the density, the survival function, and the failure rate at u, are $f_I(u) = 2\theta e^{-\theta u} - 2\theta e^{-2\theta u}$, $S_I(u) = (2e^{\theta u} - 1)e^{-2\theta u}$ and $h_I(u) = 2\theta(e^{-\theta u} - 1)/(e^{-\theta u} - 2)$, respectively. Figure 3 shows a plot of $S_C(u)$, $S_F(u)$ and $S_I(u)$ for $u \geq 0$, when $\theta = 1$ and $\delta = 0.25$. Here again, we see that system reliability under a model for cascading failures is bounded above by that under independence and below by that under causality. Clearly, for this scenario, the assumption of independence overestimates system reliability; thus it must be cautiously invoked.

A comparison of the three failure rates $h_C(u)$, $h_F(u)$ and $h_I(u)$ (see Figure 4), is instructive. It shows that $h_C(u) \to h_I(u)$, whereas $h_F(u)$ dominates the other two. Indeed, it can be verified [see Swift (2001)], that $\lim_{u\to\infty}(h_C(u) - h_I(u)) = 0$, for all values of θ and $\delta \in (0, \infty)$.

The results given above, albeit for a special case, leads us to claim that it is causal, not cascading failures, that are more deleterious to system performance, and

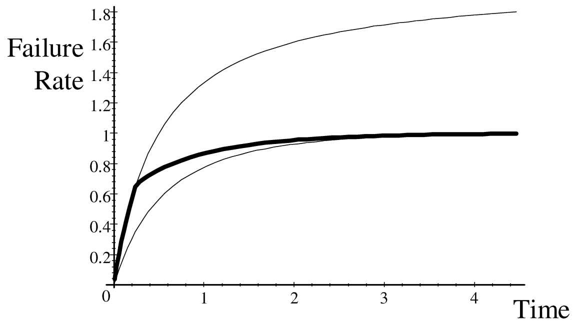

FIG. 4. *A comparison of failure rate functions under causal* (*upper line*), *cascading* (*thick line*) *and independent* (*lower line*) *failures* (*when* $\theta = 1$ *and* $\delta = 0.25$).

that an unjustified assumption of independence could result in an unwarranted overconfidence about an infrastructure's credibility. The generality of this claim is suggested by the fact that for any finite value of δ, the failure rate experiences a downward jump.

A generalization of the above model to cover monotone failure rate functions, multiple components and random values of δ, is in Swift (2001).

APPENDIX

Deterministic Causality

A deterministic view of causality has been presented via two lines of reasoning, one based on the notion of *counterfactuals*, and the other based on a principle known as *sufficiency*. The two notions do not lead to a common definition of deterministic causality.

A counterfactual is an outcome that would have been observed had the world developed differently [cf. Dawid (2000)]. As an example, consider the claim that "had Darius defeated Alexander, then Zoroastrianism would have been the dominant religion of the world." In general, if the world was in state $\mathcal{D}$, thought of as influencing $\mathcal{E}$, then we say that event $\mathcal{D}$ is the cause of an event $\mathcal{E}$, denoted $\mathcal{D} \overset{C}{\to} \mathcal{E}$, if and only if $\mathcal{E}$ does not occur in the absence of $\mathcal{D}$; for convenience, we do not recognize here multiple causes. Thus, if $\overline{\mathcal{E}}$ denotes the complement of $\mathcal{E}$, we may write, assuming that $\mathcal{P}(\overline{\mathcal{D}}) > 0$,

$$\begin{aligned}
\mathcal{D} \overset{C}{\to} \mathcal{E} \quad &\Longleftrightarrow \quad \mathcal{P}(\overline{\mathcal{E}}|\overline{\mathcal{D}}) = 1 \\
&\Longleftrightarrow \quad \mathcal{P}(\overline{\mathcal{E}} \text{ and } \overline{\mathcal{D}}) = \mathcal{P}(\overline{\mathcal{D}}) \\
&\Longleftrightarrow \quad \mathcal{E} \text{ is } \underline{\text{fully}} \text{ contained in } \mathcal{D}.
\end{aligned}$$

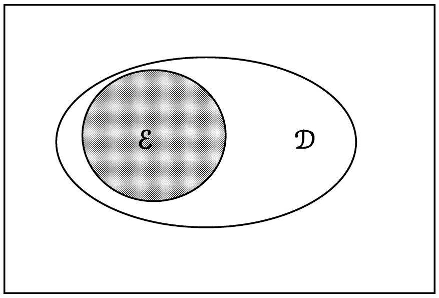

FIG. 5. *$\mathcal{D}$ is a deterministic cause of $\mathcal{E}$.*

This is illustrated in Figure 5.

However, if $\mathcal{E}$ is completely contained in $\mathcal{D}$, then

$$\mathcal{P}(\mathcal{E}|\mathcal{D}) = \frac{\mathcal{P}(\mathcal{E} \text{ and } \mathcal{D})}{\mathcal{P}(\mathcal{D})} = \frac{\mathcal{P}(\mathcal{E})}{\mathcal{P}(\mathcal{D})},$$

so that

$$\mathcal{P}(\mathcal{E}|\mathcal{D})\mathcal{P}(\mathcal{D}) = \mathcal{P}(\mathcal{E})$$

or

$$\mathcal{P}(\mathcal{E}|\mathcal{D}) \geq \mathcal{P}(\mathcal{E}),$$

and since $\mathcal{P}(\mathcal{D}) < 1$,

$$\mathcal{P}(\mathcal{E}|\mathcal{D}) > \mathcal{P}(\mathcal{E}).$$

Thus to summarize,

$$\begin{aligned}
\mathcal{D} \overset{C}{\to} \mathcal{E} \quad &\Longleftrightarrow \quad \mathcal{P}(\overline{\mathcal{E}}|\overline{\mathcal{D}}) = 1 \text{ and also that} \\
&\Longrightarrow \quad \mathcal{P}(\mathcal{E}|\mathcal{D}) > \mathcal{P}(\mathcal{E}).
\end{aligned}$$

Figure 6 depicts the three possible scenarios wherein $\mathcal{D} \overset{C}{\nrightarrow} \mathcal{E}$; that is, $\mathcal{D}$ cannot be a deterministic cause of $\mathcal{E}$.

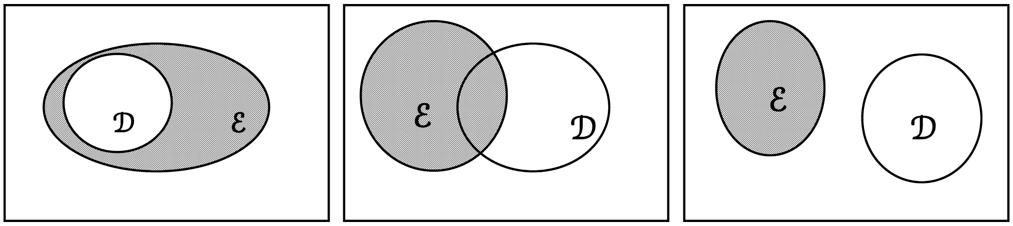

FIG. 6. *Scenarios depicting the case when $\mathcal{D}$ is not a deterministic cause of $\mathcal{E}$.*

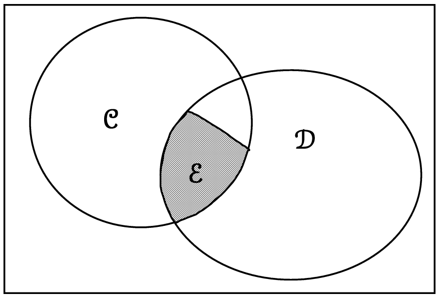

FIG. 7. *$\mathcal{C}$ and $\mathcal{D}$ are jointly a deterministic cause of $\mathcal{E}$.*

Finally, if $\mathcal{C}$ and $\mathcal{D}$ are jointly a deterministic cause of $\mathcal{E}$, then $\mathcal{E}$ must be completely contained in $\mathcal{C} \cap \mathcal{D}$, as shown in Figure 7.

The sufficiency principle for deterministic causality is based on the notion that a cause is either sufficient, or part of a sufficient condition, for the effect. This belief, common among scientists, does not mean that every event has a sufficient cause; rather, it means that if an event has a cause, then it has a sufficient cause. Thus we may say that event $\mathcal{D}$ is a sufficient cause for event $\mathcal{E}$, denoted $\mathcal{D} \overset{S}{\to} \mathcal{E}$, if $\mathcal{P}(\mathcal{E}|\mathcal{D}) = 1$, and it is easy to verify that here $\mathcal{D}$ is fully contained in $\mathcal{E}$; also, $\mathcal{P}(\mathcal{E}|\mathcal{D}) > \mathcal{P}(\mathcal{E})$, assuming that $\mathcal{P}(\mathcal{E}) < 1$. Since $\mathcal{D}$ is contained in $\mathcal{E}$ the first illustration of Figure 6 replaces Figure 5. Consequently, deterministic causality defined under counterfactuals does not imply the one defined under the sufficiency principle, and vice versa, unless $\mathcal{D}$ and $\mathcal{E}$ in Figure 5 are identical. We label the above forms of causality deterministic, since the defining probabilities, namely $\mathcal{P}(\overline{\mathcal{E}}|\overline{\mathcal{D}})$ and $\mathcal{P}(\mathcal{E}|\mathcal{D})$, are set to one.

ACKNOWLEDGMENTS

We thank Andrew Swift for his invaluable contributions to the material of Section 4.2, and the editorial team of *Statistical Science* for pointing out several issues that helped us give several finishing touches to the paper.

This paper has been written as a festschrift in honor of Professor Richard E. Barlow's birthday. Dick has made signal contributions to the theory and the applications of reliability.

This research was supported by Subcontract 35352-6085 with Cornell University under W08333-04 from the Research Institute and Grant DAAG55-97-1-0323 with the US Army Research Office.

REFERENCES

BARLOW, E. and PROSCHAN, F. (1975). *Statistical Theory of Reliability and Life Testing*. Holt, Rinehart and Winston, New York.

BERNARDO, J. and SMITH, A. F. M. (1994). *Bayesian Theory*. Wiley, New York.

COX, D. R. (1986). Comment on "Statistics and causal inference," by P. W. Holland. *J. Amer. Statist. Assoc.* **81** 963–964.

COX, D. R. (1992). Causality: Some statistical aspects. *J. Roy. Statist. Soc. Ser. A* **155** 291–301.

DAWID, A. P. (2000). Causal inference without counterfactuals. *J. Amer. Statist. Assoc.* **95** 407–424.

DIACONIS, P. and FREEDMAN, D. (1980). de Finetti's generalization of exchangeability. In *Studies in Inductive Logic and Probability II* (R. C. Jeffrey, ed.) 233–250. Univ. California Press, Berkeley.

DE FINETTI, B. (1938). Sur la condition d'equivalence partielle. In *Studies in Inductive Logic and Probability II* (1980) (R. C. Jeffrey, ed.) 193–206. Univ. California Press, Berkeley.

FREUND, J. E. (1961). A bivariate extension of the exponential distribution. *J. Amer. Statist. Assoc.* **56** 971–977.

GRANGER, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37** 424–438.

HENLEY, E. J. and KUMAMOTO, H. (1981). *Reliability Engineering and Risk Assessment*. Prentice-Hall, Englewood Cliffs, NJ.

HESSLOW, G. (1976). Two notes on the probabilistic approach to causality. *Philos. Sci.* **43** 290–292.

HESSLOW, G. (1981). Causality and determinism. *Philos. Sci.* **48** 591–605.

HOLLAND, P. W. (1986). Statistics and causal inference (with discussion). *J. Amer. Statist. Assoc.* **81** 945–970.

HUME, D. (1748). *An Inquiry Concerning Human Understanding*.

KRAJEWSKI, W. (1982). Four conceptions of causation. In *Polish Essays in the Philosophy of the Natural Sciences* (W. Krajewski, ed.) 223–235. Reidel, Dordrecht.

KULLBACK, S. and LEIBLER, R. A. (1951). On information and sufficiency. *Ann. Math. Statist.* **22** 79–86.

LINDLEY, D. V. and PHILLIPS, L. D. (1976). Inference for a Bernoulli process (a Bayesian view). *Amer. Statist.* **30** 112–119.

MARSHALL, A. W. and OLKIN, I. (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.* **62** 30–44.

MARTZ, H. F. and WALLER, R. A. (1982). *Bayesian Reliability Analysis*. Wiley, New York.

MILL, J. S. (1843). *A System of Logic*. Parker, London.

OAKES, D. (2001). *Biometrika* Centenary: Survival analysis. *Biometrika* **88** 99–142.

PEARL, J. (2000). *Causality*: *Models, Reasoning and Inference*. Cambridge Univ. Press.

ROSEN, D. A. (1978). In defense of a probabilistic theory of causality. *Philos. Sci.* **45** 604–613.

SALMON, W. C. (1980). Probabilistic causality. *Pacific Philos. Quarterly*. Univ. Southern California.

SUPPES, P. (1970). *A Probabilistic Theory of Causality*. North-Holland, Amsterdam.

SWIFT, A. (2001). Stochastic models of cascading failures. Ph.D. dissertation, George Washington Univ.

VON MISES, R. (1957). *Probability, Statistics and Truth*. Allen and Unwin, London.