The Impact of Lehmann's Work in Applied Probability

Walter W. Piegorsch and Moshe Shaked

Lehmann's works in applied probability crossed a broad spectrum of topics, and many of his papers had substantial impact. We comment on two such examples below. The first [19] provided a valuable beginning for the study and application of notions of positive and negative dependence. The second [20] also touched on issues of probability theory and the nature of random variables—in particular, on the distributional aspects of the reciprocal of a random variable. The contributions each represented are a study in contrasts, however: the former was a substantial work from which many wide-ranging contributions developed, while the second was a shorter, later-appearing piece whose impact(s) are still evolving.

1 Comments on Lehmann [19]

Erich Lehmann introduced and studied in [19] three concepts of positive dependence of two random variables. These are the concepts of positive quadrant dependence (PQD), positive regression dependence (PRD), and positive likelihood ratio dependence (PLRD). In the same paper Lehmann also introduced and studied the negative dependence analogs of these concepts. Since the publication of [19], these concepts have been used and applied in diverse areas of applied probability and statistics. They can be found in some abstract probabilistic papers, as well as in practical studies in medicine, operations research, management science, biology, sociology, economics, and physics. For example, a search in the ISI Web of Knowledge on May 11, 2009, yielded more than 500 papers that cite [19]. Apart from mainstream journals in statistics and probability, the references to [19] included papers in journals such as *Monthly Weather Review*, *Control and Cybernetics*, *Pharmaceutical Statistics*,

W.W. Piegorsch and M. Shaked

Department of Mathematics, University of Arizona, Tucson, Arizona 85721, USA
e-mail: piegorsch@math.arizona.edu; shaked@math.arizona.edu

Journal of Hydrologic Engineering, Stochastic Hydrology and Hydraulics, Journal of Geophysical Research, Progress in Natural Science, Psychometrika, Multivariate Behavioral Research, Journal of Mathematical Psychology, Applied Psychological Measurement, and Stochastic Environmental Research and Risk Assessment, to name a few.

The study in [19] was a main starting point for later studies of concepts of stochastic positive and negative dependence. These studies continue up to the present time. Two main routes leading from [19] were used by researchers in the area of dependence concepts. One type consists of studies that introduce further notions of *bivariate* dependence. For example, Harris [10] studied the notion of right corner set increasing dependence (that, in fact, was already mentioned in [19], but was not thoroughly studied there). Yanagimoto [31] refined the notions of Lehmann [19] and of Harris [10], and thus introduced some families of positive dependence notions. A more systematic extension of all the above bivariate notions of dependence was made in Shaked [24]. Another type involves studies that extend the concepts of [19] to the *multivariate* case. A basic paper in this vein is the paper by Esary, Proschan, and Walkup [9] in which the notion of association of a set (not necessarily a pair) of random variables is introduced. The multivariate ideas in [9], that built on [19], led to a large number of further studies. Perhaps it is worthwhile to mention here the work of Lee [17] which extended the positive dependence notions of Shaked [24] from the bivariate to the multivariate case. In another paper of Lee [18] a parallel development of negative dependence notions is given. Other studies of positive and negative dependence concepts, that start from [19] and extend it in various directions, can be found in the papers by Block, Savits, and Shaked [3,4].

It is intriguing to speculate how familiar Lehmann was with the relationship between some of his concepts of positive dependence in [19] and the terminology and basic results of the theory of total positivity; see Karlin [12]. For example, the notion of PLRD is actually the same as requiring the joint density function of the two underlying random variables to be totally positive of order 2 (TP_2). By noticing the total positivity condition above, the relation of the notions of Lehmann [19] to notions in statistical physics become apparent, and new routes open for their extensions to the multivariate case. For example, the PLRD notion can be extended to the multivariate case by requiring the joint density function of the underlying random variables to be TP_2 in pairs. Another multivariate extension of the PLRD notion is obtained by requiring the joint density function of the underlying random variables to satisfy another condition that is an extension of the bivariate TP_2 condition; see, for example, Karlin and Rinott [13]. Block and Ting [5] studied these kinds of extensions.

The three main concepts of positive dependence in [19] are related as follows

$$\text{PLRD} \implies \text{PRD} \implies \text{PQD}. \quad (1.1)$$

It follows that every property that PQD random vectors have, is also a property of PLRD and PRD random vectors. For example, Lehmann [19] proved that if the covariance of a PQD (and hence also of a PRD or a PLRD) random vector exists then

it must be nonnegative. Lehmann [19] also showed that other measures of association, such as Kendall's τ , Spearman's ρ , and Blomqvist's q , are all nonnegative for PQD random vectors.

An important and useful closure property of the PQD concepts that is proven in [19] is the following. Suppose that the independent random vectors $(Y_1, V_1), (Y_2, V_2), \dots, (Y_n, V_n)$ are all PQD. Let r and s be n -dimensional real functions that are nondecreasing in each of their n variables when the other variables are held fixed. Then

$$r(Y_1, Y_2, \dots, Y_n) \quad \text{and} \quad s(V_1, V_2, \dots, V_n)$$

are PQD. In particular the PQD concept is closed under nondecreasing transformations of the underlying random variables.

By looking at the formal definition of the PQD concept in [19] (which is also given in (1.2) below) it is easy to see that for every random variable X we have that (X, X) is PQD. From the closure property that is described in the previous paragraph it follows that $(X_1, X_2) \stackrel{\text{def}}{=} (X, s(X))$ is PQD whenever s is a nondecreasing function. The random variables X_1 and X_2 described above are called comonotone in the recent actuary and insurance literature; see, for instance, Ludkovski and Rüschendorf [21] and references therein. From the above discussion it follows that comonotone random variables are positively dependent in the sense of all the notions in [19].

An elegant and useful result that is proven in [19] shows that if a PQD random vector (X_1, X_2) has a zero covariance then X_1 and X_2 must be independent. This property has been known for bivariate normal random variables – the above result extends this surprising property to a host of jointly distributed pairs of random variables.

The concepts of positive and negative dependence in [19] (and also their later extensions) have been applied in many areas of statistics and probability. Lehmann himself, in [19], applied them to some hypotheses testing situations. One application involves a multiple comparisons setting. Suppose that we have a number of pairs of hypotheses that are tested using the same sample. Then the test statistics are not independent, and it is often of interest to control not the individual significance levels (that is, error probabilities), but the experimentwise error rate. Lehmann [19] noticed that sometimes the test statistics are positively or negatively dependent, and as a consequence he obtained tight bounds on the experimentwise error rate.

Another application of the concepts of positive dependence in statistics, that Lehmann described in [19], is concerned with testing the null hypothesis H_0 of independence versus the alternative H_a of positive dependence. One benefit of the various definitions of positive dependence in [19] is that they enable the statistician to formally state H_a . Furthermore, by identifying some quantity that must be positive under H_a (that is, under the assumption of positive dependence), but 0 under H_0 , one can use the empirical sample estimator of that quantity as a test statistic. Lehmann [19] showed that such testing procedures yield unbiased tests. This very idea has been used extensively again and again in later literature. For example, see Shetty and Pandit [27] and references therein.

Other applications of the concepts of positive and negative dependence are based on various inequalities involving useful probabilistic quantities that follow from these

concepts. For example, Barlow and Proschan [1] obtained useful bounds on reliability systems using the notions of [19], whereas various useful inequalities in queueing theory have been derived using the theory of positive dependence; see, for instance, Daduna and Szekli [7] and references therein.

Positive or negative dependence concepts are a useful and convenient tool for relaxing the assumption of independence in various studies of limit theorems. Many papers have adopted this tool. We cannot list them all here, but we will just mention a recent paper, by Ko, Kim, and Han [16], in which the reader may find further references for the use of positive and negative dependence in obtaining limit theorems.

For the purpose of motivating the following discussion, recall formally from [19] that the two random variables X_1 and X_2 are PQD if

$$P\{X_1 \leq x_1, X_2 \leq x_2\} \geq P\{X_1 \leq x_1\}P\{X_2 \leq x_2\} \quad \text{for all } (x_1, x_2) \in \mathbb{R}^2. \quad (1.2)$$

Notice that the inequality (1.2) compares the probability that (X_1, X_2) is in a lower quadrant, with the probability of being in the same lower quadrant if X_1 and X_2 “were independent”. This observation led several researchers to introduce and consider a stochastic order between bivariate random vectors with the same marginals, (X_1, X_2) and (Y_1, Y_2) , say, through the definition

$$P\{X_1 \leq x_1, X_2 \leq x_2\} \leq P\{Y_1 \leq x_1, Y_2 \leq x_2\} \quad \text{for all } (x_1, x_2) \in \mathbb{R}^2; \quad (1.3)$$

we may denote (1.3) as $(X_1, X_2) \leq_{\text{PQD}} (Y_1, Y_2)$, and call it the PQD positive dependence stochastic order. Intuitively, $(X_1, X_2) \leq_{\text{PQD}} (Y_1, Y_2)$ indicates that Y_1 and Y_2 are “more positively dependent” than X_1 and X_2 . Tchen [28], among others, studied this order.

For the purpose of comparing multivariate (rather than just bivariate) random vectors in the sense of positive dependence, one may consider extending the defining inequality (1.3) to the multivariate case. It turns out that there is more than one way of doing this. Some authors considered defining a positive dependence comparison of random vectors by comparing probabilities of lower or upper orthants in $\mathbb{R}^n$; see, for example, Joe [11]. Others, such as Bäuerle [2] and Shaked and Shanthikumar [25], studied a stronger integral stochastic order, which compares the expectations of supermodular functions of the compared random vectors.

The above discussion illustrates how the PQD concept in [19] led to the introduction and study of corresponding positive dependence stochastic orders. A similar discussion applies to the other two positive dependence concepts in (1.1). For example, a stochastic order that is based on the PRD notion has been introduced and studied in Yanagimoto and Sibuya [32], whereas a stochastic order that is based on the PLRD notion was considered in Kimeldorf and Sampson [15]. If we denote the corresponding orders by $\leq_{\text{PRD}}$ and $\leq_{\text{PLRD}}$, then, motivated by (1.1), we should expect the following relationships among the three stochastic orders above:

$$\leq_{\text{PLRD}} \implies \leq_{\text{PRD}} \implies \leq_{\text{PQD}} .$$

Generally speaking, this is indeed the case, but the interested reader is advised to consult Chapter 9 in Shaked and Shanthikumar [26] where a detailed study of this issue is given.

2 Comments on Lehmann and Shaffer [20]

In contrast to the far-reaching contributions made by Lehmann's work [19] on positive dependence, the article [20] by Lehmann and Shaffer on inverted distributions has to date seen lesser development. The two papers do share a common feature, however: their primary focus is on distributional attributes of random variables rather than new statistical methodology. This leads us to speculate that the future, potential impact of [20] is still untapped.

The questions dealt with in this article [20] are, on the surface, somewhat straightforward: how can the relationship between the probability density function f_X of a continuous random variable X , and that of its inverted distribution—the probability density function f_Y of $Y = 1/X$ —be characterized? Lehmann and Shaffer [20] motivate the question from previous discussions on the nature of the inverse moment $E[1/X]$ (also called the negative moment or the reciprocal moment by some authors; see, e.g., Khuri and Casella [14], Piegorsch and Casella [23], Wooff [30], or Cressie, Davis, Folks, and Policello [6]). Indeed, their discussion concerns only continuous random variables, since clearly, if a discrete probability mass function has non-zero mass at $X = 0$, $E[1/X]$ does not exist. This makes the issue of inverted distributions for discrete random variables somewhat less interesting.

Expanding on results in [23], Lehmann and Shaffer [20] observe that the nature of $f_X(x)$ near $x = 0$ helps characterize the nature of the inverse probability density function f_Y , and further that the behavior of $f_X(x)$ as $x \rightarrow \infty$ can tell us something about the behavior of $f_Y(y)$ near $y = 0$. This is not surprising—as the authors themselves note [20, page 192], “it is difficult to imagine that these results are new, but we have not been able to find them in the literature.” Indeed, the basic mathematical finding simply reflects the relationship between the behavior of $f_X(x)$ near $x = 0$ and the stability of the integral $\int x^{-1} f_X(x) dx$. Nonetheless, some fascinating distributional results emerge from the effort. Lehmann and Shaffer [20] show that if

$$x^2 f_X(x) \rightarrow 0 \text{ as } |x| \rightarrow \infty, \quad (2.4)$$

then $f_Y(y) \rightarrow 0$ as $y \rightarrow 0$. To do so, they characterize how heavy the tail behavior of f_X appears, related to what the authors call *Cauchy tails*. Specifically, they find that if $f_X(0+) = \lim_{x \rightarrow 0+} f_X(x)$, then the ratio of $f_Y(y)$ to $1/(1+y^2)$ tends to $f_X(0+)$ as $y \rightarrow \infty$. Thus if $f_X(0+) = 0$ the (right) tail behavior of f_Y is lighter than that of a Cauchy probability density function, if $f_X(0+)$ is positive and finite the tail behavior of f_Y is similar to the Cauchy, and if $f_X(0+)$ is infinite the tail behavior is heavier than Cauchy. (Analogous left-tail results are obtained by studying $x \rightarrow 0-$ and $y \rightarrow -\infty$.)

Lehmann and Shaffer [20] also use these tail-behavior results to describe the modal nature of f_Y . In particular, and rather intriguingly, they elicit that if (2.4) holds such that $f_X(x) > 0$ for both positive and negative values of x , then f_Y will be multi-modal, with at least one mode at some value of $y < 0$ and one mode at some value of $y > 0$. (In most cases, the inverted probability density function is exactly bimodal.) They expand upon these results via some clever limiting operations, showing, e.g., that the strange bi- or multi-modal behavior of an appropriately standardized version of $1/X$ can actually disappear as certain location and scale parameters are driven to their extremes. (See Sections 3 and 4 of the Lehmann and Shaffer [20] for more details.) The authors refer to this spectrum of behavior for $1/X$ as “bizarre,” a term difficult to argue with, but in doing so they also yield insights into the fascinating and sometimes unusual nature of inverted distributions.

As we note above, the impact of the Lehmann and Shaffer [20] article has been limited, perhaps due to its focus on these curious aspects of (inverted) distributional theory. A few applications have appeared for selected econometric problems—see, e.g., Diebold and Lamb [8] and ten Raa and Steel [29]—but these have generally been confined to calculations on inverse moments and not to the more general features of the inverted distribution. An interesting, hierarchical modeling effort is given by Peng [22], who estimates an inverse moment via Bayesian methods in an engineering reliability analysis. Indeed, we speculate that when working with a posterior probability density function $\pi(\mu|X)$, characterization of the larger distributional features of a reciprocal parameter such as $\psi = 1/\mu$, beyond just the inverse moment $E[\psi|X]$, may have unexploited value. The Lehmann and Shaffer [20] results available for the tail and modal behavior of an inverted (posterior) distribution $\pi(\psi|X)$ might be important considerations in future Bayesian studies with inverse or reciprocal quantities such as $\psi = 1/\mu$.

References

- [1] Barlow, R. E. and Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing, Probability Models*. Holt, Rinehart, and Winston, New York.
- [2] Bäuerle, N. (1997). Inequalities for stochastic models via supermodular orderings. *Communications in Statistics - Stochastic Models* **13**, 181–201.
- [3] Block, H. W., Savits, T. H., and Shaked, M. (1982). Some concepts of negative dependence. *Annals of Probability* **10**, 765–772.
- [4] Block, H. W., Savits, T. H., and Shaked, M. (1985). A concept of negative dependence using stochastic ordering. *Statistics and Probability Letters* **3**, 81–86.
- [5] Block, H. W. and Ting, M.-L. (1981). Some concepts of multivariate dependence. *Communications in Statistics A - Theory and Methods* **10**, 749–762.
- [6] Cressie, N., Davis, A. S., Folks, J. L., and Policello, G. E., II (1981). The moment-generating function and negative integer moments. *American Statistician* **35**, 148–150.
- [7] Daduna, H. and Szekli, R. (2004). On the correlation structure of closed queueing networks. *Stochastic Models* **20**, 1–29.
- [8] Diebold, F. X. and Lamb, R. L. (1997). Why are estimates of agricultural supply response so variable? *Journal of Econometrics* **76**, 357–373.

- [9] Esary, J. D., Proschan, F., and Walkup, D. W. (1967). Association of random variables, with applications. *Annals of Mathematical Statistics* **38**, 1466–1474.
- [10] Harris, R. (1970). A multivariate definition for increasing hazard rate distributions. *Annals of Mathematical Statistics* **41**, 713–717.
- [11] Joe, H. (1997). *Multivariate Models and Dependence Concepts*, Chapman and Hall, London.
- [12] Karlin, S. (1968). *Total Positivity*, Stanford University Press, Stanford, California.
- [13] Karlin, S. and Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *Journal of Multivariate Analysis* **10**, 467–498.
- [14] Khuri, A. and Casella, G. (2002). The existence of the first negative moment revisited. *American Statistician* **56**, 44–47.
- [15] Kimeldorf, G. and Sampson, A. R. (1987). Positive dependence orderings. *Annals of the Institute of Statistical Mathematics* **39**, 113–128.
- [16] Ko, M.-H., Kim, T.-S., and Han, K.-H. (2009). A note on the almost sure convergence for dependent random variables in a Hilbert space. *Journal of Theoretical Probability* **22**, 506–513.
- [17] Lee, M.-L. T. (1985). Dependence by total positivity. *Annals of Probability* **13**, 572–582.
- [18] Lee, M.-L. T. (1985). Dependence by reverse regular rule. *Annals of Probability* **13**, 583–591.
- [19] Lehmann, E. L. (1966). Some concepts of dependence. *Annals of Mathematical Statistics* **37**, 1137–1153.
- [20] Lehmann, E. L. and Shaffer, J. P. (1988). Inverted distributions. *American Statistician* **42**, 191–194.
- [21] Ludkovski, M. and Rüschendorf, L. (2008). On comonotonicity of Pareto optimal risk sharing. *Statistics and Probability Letters* **78**, 1181–1188.
- [22] Peng, C. Y. (2008). The first negative moment in the sense of the Cauchy principal value. *Statistics and Probability Letters* **78**, 1765–1774.
- [23] Piegorsch, W. W. and Casella, G. (1985). The existence of the first negative moment. *American Statistician* **39**, 60–62.
- [24] Shaked, M. (1977). A family of concepts of dependence for bivariate distributions. *Journal of the American Statistical Association* **72**, 642–650.
- [25] Shaked, M. and Shanthikumar, J. G. (1997). Supermodular stochastic orders and positive dependence of random vectors. *Journal of Multivariate Analysis* **61**, 86–101.
- [26] Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic Orders*, Springer, New York.
- [27] Shetty, I. D. and Pandit, P. V. (2003). Distribution-free tests for independence against positive quadrant dependence: A generalization. *Statistical Methods and Applications* **12**, 5–17.
- [28] Tchen, A. H. (1980). Inequalities for distributions with given marginals. *Annals of Probability* **8**, 814–827.
- [29] ten Raa, T. and Steel, M. F. J. (1994). Revised stochastic-analysis of an input-output model. *Regional Science and Urban Economics* **24**, 361–371.
- [30] Wooff, D. A. (1985). Bounds on reciprocal moments with applications and developments in Stein estimation and post-stratification. *Journal of the Royal Statistical Society, Series B* **47**, 362–371.
- [31] Yanagimoto, T. (1972). Families of positively dependent random variables. *Annals of the Institute of Statistical mathematics* **24**, 559–573.
- [32] Yanagimoto, T. and Sibuya, M. (1972). Stochastically larger component of a random vector. *Annals of the Institute of Statistical Mathematics* **24**, 259–269.