THE CONVEX ANALYSIS OF RANDOM VARIABLES

R. Tyrrell Rockafellar[1]

Abstract

Any real-valued random variable induces a probability distribution on the real line which can be described by a cumulative distribution function. When the vertical gaps that may occur in the graph of that function are filled in, one gets a maximal monotone relation which describes the random variable by its characteristic curve. Maximal monotone relations in the plane are known in convex analysis to correspond to the subdifferentials of the closed proper convex functions on the real line. Here that connection is developed in terms of what those convex functions and their conjugates say about the random variables, and how that information serves in applications to stochastic optimization.

Keywords: *convex analysis, random variables, cumulative distribution functions, characteristic curves, quantiles, superquantiles, VaR, CVaR, stochastic optimization*

[1]University of Washington, Department of Mathematics, Box 354350, Seattle, WA 98195-4350;
E-mail: *rtr@uw.edu*, URL: http://www.math.washington.edu/∼rtr/mypage.html

1 Distribution Functions and Characteristic Curves

Random variables are central to statistics but vital also in areas of applied mathematics that deal with stochastic modeling, such as engineering, finance and operations research. Less obvious, however, and hardly known, are some deep connections with convex analysis. The purpose of this note is to explain those connections and their implications for computational work. Further details, references, and proofs of various assertions are provided in [9].

Any random variable X (real-valued) induces a probability distribution on the real line $I\!R$ through the rule that the probability $\text{prob}_X\{I\}$ of an interval $I \subset I\!R$ is the probability of X having outcome in I. Two random variables X and Y can of course induce the same probability distribution on $I\!R$, $\text{prob}_X = \text{prob}_Y$, without X and Y being identical, because this aspect of them does not reflect the degree to which they may be independent or correlated. Nonetheless, the induced distributions on $I\!R$ are helpful in understanding many features of randomness.

A convenient handle on the probability distribution prob_X is the *cumulative distribution function* F_X, defined by

$$F_X(q) = \text{prob}_X\{\,(-\infty, q]\,\} \;\text{ for }\; q \in I\!R. \tag{1}$$

This is a nondecreasing function from $I\!R$ into $[0,1]$ which is right continuous and tends to 1 as $x \nearrow \infty$ and to 0 as $x \searrow -\infty$. It completely determines prob_X. On the other hand, any function from $I\!R$ to $[0,1]$ with the listed properties is prob_X for some random variable X (far from unique).

Although most of what follows is concerned only with probability distributions on $I\!R$ in themselves, apart from a particular X, we maintain X in our notation as a reminder of the context in which the ideas are to be applied.

Tradition in statistics looks to F_X, but it is equally possible to characterize prob_X in terms of the function

$$F_X^-(q) = \text{prob}_X\,\{\,(-\infty, q)\} \;\text{ for }\; q \in I\!R, \tag{2}$$

which has the same properties except for being *left* continuous instead of *right* continuous. Indeed, F_X and F_X^- are locked together by the limit relations

$$F_X^-(q) = \lim_{q' \nearrow q} F_X(q'), \qquad F_X(q) = \lim_{q' \searrow q} F_X^-(q'). \tag{3}$$

When $F_X = F_X^-$, the random variable X is said to be *continuous*, but the extreme opposite case, where these functions agree except at finitely many points where a "jump" occurs, and are constant between them, is likewise valuable in corresponding to a random variable X that is *discrete*.

Despite the fact that standard theory can, and ordinarily does, make do with just F_X and ignore F_X^-, there is a key reason to pay attention to both and even to go a step further in considering the set

$$\Gamma_X = \{\,(q,p)\,|\, F_X^- \le p \le F_X(q)\} \subset I\!R \times I\!R$$

as the *characteristic curve* associated with X. In geometric terms, Γ_X is obtained from the graph of F_X (or for that matter F_X^-) simply by filling in the vertical segments wherever there is a gap caused by a jump. It, too, completely characterizes the probability distribution prob_X, but what advantage does it hold? The advantage is that Γ_X directly opens the door to convex analysis by being a *maximal monotone relation* in $I\!R \times I\!R$.

Much about maximal monotone relations in $I\!R \times I\!R$ can be found in [6] and [15] The topic was first studied by G. Minty [3] in the framework of variational principles (i.e., optimization rules) that characterize equilibrium configurations of flows and potentials in general networks — electrical, hydraulic, economic and much more. It was taken up soon afterward in the PhD thesis [5] and subsequently expanded in the book [7].

A subset Γ of $I\!R \times I\!R$ is said to give a *monotone relation* when

$$(q,p) \in \Gamma,\ (q',p') \in \Gamma \quad \Longrightarrow \quad (q'-q)\cdot(p'-p) \geq 0, \tag{4}$$

or equivalently

$$(q,p) \in \Gamma,\ (q',p') \in \Gamma \quad \Longrightarrow \quad \text{either } (q',p') \geq (q,p) \text{ or } (q,p) \geq (q',p') \tag{5}$$

in the coordinatewise partial ordering of $I\!R \times I\!R$. (The monotonicity concept in convex analysis extends to subsets of $I\!R^n \times I\!R^n$ through (4), but without equivalence to (5); see [6] and [15].)

A monotone relation Γ in $I\!R \times I\!R$ is *maximal* if there does not exist a monotone relation Γ' in $I\!R \times I\!R$ with $\Gamma' \supset \Gamma$ and $\Gamma' \neq \Gamma$. Any monotone relation can be extended to one that is maximal. The *domain* and *range* of a maximal monotone relation, namely the projections

$$\operatorname{dom}\Gamma = \{\, q \,|\, \exists p \text{ with } (q,p) \in \Gamma\}, \qquad \operatorname{rge}\Gamma = \{\, p \,|\, \exists q \text{ with } (q,p) \in \Gamma\}, \tag{6}$$

are always *intervals* in $I\!R$. In our context of random variables X and their characteristic curves Γ_X, we are dealing with the class of maximal monotone relations Γ such that

$$\operatorname{dom}\Gamma = (-\infty,\infty), \qquad (0,1) \subset \operatorname{rge}\Gamma \subset [0,1]. \tag{7}$$

The main fact about maximal monotone relations $\Gamma \subset I\!R \times I\!R$ from the perspective of convex analysis is that they are the *graphs of the subdifferentials ∂f of the closed proper convex functions f on $I\!R$*. Specifically, Γ is maximal monotone if and only if there is a lower semicontinuous convex function $f : I\!R \to (-\infty,\infty]$, $f \not\equiv \infty$, such that

$$\Gamma = \{\,(q,p) \,|\, p \in \partial f(q)\}, \tag{8}$$

where

$$\partial f(q) = \{\, p \,|\, f(q') \geq f(q) + p\cdot(q'-q)\ \forall q'\}, \tag{9}$$

or in terms of the left and right derivatives f'^{-} and f'^{+},

$$\partial f(q) = \{\, p \,|\, f'^{-}(q) \leq p \leq f'^{+}(q)\,\} \tag{10}$$

(under the convention that $f'^{-}(q) = -\infty$ at the left endpoint of $\operatorname{dom} f$ and below it, whereas $f'^{+}(q) = \infty$ at the right endpoint of $\operatorname{dom} f$ and above it). The function f is uniquely determined by Γ up to an additive constant.

What might this mean for random variables? In that case, since $\operatorname{dom}\Gamma_X = I\!R$, f is a *finite* convex function on $I\!R$ having F_X^- and F_X as its left and right derivatives f'^{-} and f'^{+}.

Statistics is an old subject and convex analysis is getting to be one, so why has this connection between random variables and convex functions been largely out of sight in the literature? Actually, in one way, it has long been present, even if unexploited, through the topic of stochastic dominance.

A random variable X is said to exhibit *first-order stochastic dominance* over a random variable Y, in the context of higher outcomes being preferred to lower outcomes, if $F_X \leq F_Y$, or in other words, if for every $q \in I\!R$ the probability of X being $\leq q$ is no greater than the probability of Y being $\leq q$. In the same context, *second-order stochastic dominance* corresponds instead to having $F_X^{(2)} \geq F_Y^{(2)}$, where

$$F_X^{(2)}(q) = \int_{-\infty}^{q} F_X(q')dq'. \tag{11}$$

This integral is sure to be finite as long as $E[\,|X|\,] < \infty$, which is *a restiction we henceforth place on all the random variables in this note.*

Clearly, because F_X is nondecreasing, the function $F_X^{(2)}$ is indeed convex and has F_X^- and F_X as its left and right derivatives f'^- and f'^+. Furthermore it is nondecreasing and tends to 0 as $q \searrow -\infty$ and to ∞ as $q \nearrow \infty$. However, for some purposes, especially in the opposite context where the random variable X represents a "cost", so that lower outcomes are preferable to higher outcomes, it may not be the ideal choice for the convex function associated with X; an adjustment in the additive constant of integration may be desirable. In that connection our focus will instead be on the *superexpectation* function E_X, defined by

$$E_X(q) = E[\max\{X, q\}], \tag{12}$$

which likewise is convex and nondecreasing with the same left and right derivatives, but instead tends to $E[X]$ as $q \searrow -\infty$. This vertical shift from $F_X^{(2)}$ may seem inconsequential, but the fact that the size of the shift depends on X rather than being fixed in advance has a significant advantage. We will return to this close to the end of this note, after other ingredients have been added to the conceptual mix.

2 Quantiles and Superquantiles

The distribution function F_X is not the only important function associated with a random variable X. Another is the *quantile function* Q_X, given by

$$\begin{aligned} Q_X(p) &= \min\{\, q \,|\, F_X(q) \geq p\} \ \text{ for } \ p \in [0,1) \\ &= \text{lowest } p \text{ such that } \ \text{prob}\{\, X > q\} \ \leq \ 1-p. \end{aligned} \tag{13}$$

This is a left-continuous nondecreasing function having $Q_X(0)$ as the essential infimum of X, possibly $-\infty$. For our purposes it is appropriately extended from $[0,1)$ to $(-\infty, \infty)$ by taking

$$Q_X(p) = -\infty \ \text{ when } p < 0, \qquad Q_X(p) = \infty \ \text{ when } p \geq 1. \tag{14}$$

The class of functions arising as Q_X for some X is then the class of all nondecreasing left-continuous functions on $I\!R$ satisfying (14). The right-continuous partner to Q_X is

$$Q_X^+(p) = \lim_{p' \searrow p} Q_X(p'). \tag{15}$$

From either Q_X or Q_X^+ the probability distribution prob_X can be recovered completely.

We can associate maximal monotone relations with quantile functions just as with distribution functions, essentially by filling in the vertical gaps in their graphs. Specifically, a maximal monotone relation Δ_X is given by

$$\Delta_X = \{\, (p,q)] \in I\!R \times I\!R \,|\, Q_X(p) \leq q \leq Q_X^+(p)\,\}. \tag{16}$$

This is inversely related to the maximal monotone relation Γ_X in the sense that

$$\Delta_X = \{\, (p,q) \,|\, (q,p) \in \Gamma_X\}, \qquad \Gamma_X = \{\, (q,p) \,|\, (p,q) \in \Delta_X\}. \tag{17}$$

Indeed, in the special case where F_X is continuous and everywhere increasing (not just nondecreasing), so that Γ_X contains neither vertical line segments nor horizontal line segments, (17) reduces to $Q_X = F_X^{-1}$ on $(0,1)$.

In this quantile context, the question of an associated convex function returns in parallel fashion. Again we have the existence of a closed, proper convex function $g : I\!R \to (-\infty, \infty]$, $g \not\equiv \infty$, which is unique up to an additive constant and yields

$$\Delta_X = \operatorname{gph} \partial g = \{ (p,q) \,|\, q \in \partial g(p) \} = \{ (p,q) \,|\, g'^{-}(p) \le q \le g'^{+}(p) \}. \tag{18}$$

Then Q_X is the left derivative g'^{-}, whereas Q_X^+ is the right derivative g'^{+}. But because of the inverse relationship between Γ_X and Δ_X in (17), convex analysis tells us further that, the free additive constant in g can be coordinated with the one for f in (8) in order to make f and g be *conjugate* convex functions in the sense of the Legendre-Fenchel transform:

$$g(p) = f^*(p) = \sup_q \{ pq - f(q) \}, \qquad f(q) = g^*(q) = \sup_p \{ pq - g(p) \}. \tag{19}$$

With f and g tied together in this way, there is still a degree of freedom; shifting f upward by some amount corresponds to shifting g downward by that amount, and vice versa. Is there a particularly advantageous way to exploit that remaining freedom?

It has already been noted that one choice of f is the function $F_X^{(2)}$ in (11). For that choice, the conjugate function f^* has been calculated by Ogryczak and Ruszczynski in [4, Theorem 3.1] (2002) to be

$$f^*(p) = F_X^{(-2)}(p) = \int_0^p Q_X(p')dp' - g(p),$$

which has the well-known Lorenz curve [2] as its graph. The alternative choice of f as the superexpectation function E_X in (12) yields instead

$$f^*(p) = E_X^*(p) = -(1-p)\overline{Q}_X(p) - g(p), \tag{20}$$

where $\overline{Q}_X$ is the *superquantile function* associated with X, expressed by

$$\overline{Q}_X(p) = \begin{cases} \frac{1}{1-p} \int_p^1 Q_X(p')dp' & \text{for } p \in (0,1), \\ 0 & \text{for } p = 1, \\ E[X] & \text{for } p = 0, \end{cases} \tag{21}$$

and extended beyond $[0,1]$ by $-\infty$ to the left and ∞ to the right, just like Q_X above. The superquantile function $\overline{Q}_X$ is continuous and increasing on $[0,1]$, and $\overline{Q}_X \ge Q_X$.

Conjugate duality in the form of (20) was already known in [11] (2000), at least implicitly, as the basis for passing between "value-at-risk" and "conditional value-at-risk" in assessments of uncertainty. That will be recounted in the next section. The "superquantial" terminology is more recent, however, having been proposed in [8] (2010) as better for applications outside of finance, such as engineering reliability.

3 Connections with Risk and Optimization Under Uncertainty

The usual focus in optimization for reasons of standardization and general exposition is on minimization, instead of maximization, and inequality constraints of type $\le$ instead of $\ge$, although all situations are ultimately covered anyway through changes of sign. In that environment, quantities to be minimized, or to be held below some upper bound, can be regarded abstractly as "costs" of some sort, with "losses," "damages," and "hazard levels" as possible linguistic substitutes.

A typical situation in finance, engineering and many other areas of application involves random values X representing such "costs," coming in the future, which can be influenced to some extent, at least, by a decision taken in the present. In viewing that decision as a vector $u = (u_1, \ldots, u_n)$, say, we can think of a "cost" random variable $X(u)$ depending on the choice of u from some set $U \subset I\!R^n$ specified by various underlying constraints.

We might like to limit the choice $u \in U$ so as to ensure that no outcomes of $X(u)$ exceed a particular upper bound b. This may be impossible to fulfill, though. Every available $u \in U$ may carry with it some circumstances that violate this, and a compromise of sorts is then essential. The critical issue is how to articulate, flexibly, a mathematical model for having

$$X(u) \;\text{“adequately”}\; \leq b. \tag{22}$$

The idea behind a *measure of risk*, as a functional $\mathcal{R}$ that assigns numerical values (possibly ∞) to random variables, is to convert (22) into an ordinary inequality involving a numerical function generated from $\mathcal{R}$, namely

$$f_{\mathcal{R}}(u) \leq b \;\text{ where }\; f_{\mathcal{R}}(u) = \mathcal{R}(X(u)), \tag{23}$$

which can then be handled by familiar methodology of optimization.

For instance, the case of $\mathcal{R}(X) = \sup X$ (the essential supremum of a random variable X) would translate (22) into having $X(u) \leq b$ with probability 1. On the other hand, the case of $\mathcal{R}(X) = E[X]$ would translate it into requiring that $X(u)$ be $\leq b$ "on average."

There are many possibilities for $\mathcal{R}$, but a particularly popular one is to choose a probability level $p \in (0, 1)$ and take $\mathcal{R}(X)$ to be the pth quantile $Q_X(p)$ of the random variable X. This is attractive because

$$Q_X(p) \leq b \iff \text{prob}\{\, X \leq b\} \geq 1 - p. \tag{24}$$

The model (23) then interprets (22) as allowing $X(u)$ to exceed b only with a probability less than $1 - p$.

Although simple and seemingly very natural, this approach has serious drawbacks. Trouble comes from the fact that the expression $f_{\mathcal{R}}(u) = Q_{X(u)}(p)$ can have poor behavior with with respect to the decision vector u on which the random cost $X(u)$ depends. It might not even be continuous in u, not to speak of being differentiable, and this can make it difficult to work with in a context of optimization. For instance, if the setting is that of discrete probability, $Q_{X(u)}$ is a step function of $p \in [0, 1]$ described by finitely many jumps, and smooth behavior with respect to shifts in u is out of question.

Another drawback to quantiles is that in interpreting (22) in terms of the inequality being satified with a specified probability there is no reflection of how bad the inevitable exceptions to the inequality may be for the situation being modeled.

A better approach may be to replace quantiles $Q_X(p)$ by superquantiles $\overline{Q}_X(p)$. The superquantile version of (24) is

$$\overline{Q}_X(p) \leq b \iff \begin{cases} \text{even in the upper } p\text{-tail of the probability distribution} \\ \text{prob}_X \text{ on } (-\infty, \infty)\text{, one has outcomes } X \leq b \text{ “on average.”} \end{cases} \tag{25}$$

The p-tail distribution associated with X refers to the probability distribution obtained by passing from the distribution function F_X to the truncated distribution function

$$F_X^p(q) = \frac{1}{1-p} \max\{0, F_X(q) - p\}. \tag{26}$$

When F_X has no jump at $q = Q_X(p)$, this gives the conditional probability distribution induced by prob_X on the interval $[Q_X(p), \infty)$, and then $\overline{Q}_X(p)$ is the expectation of X conditional on having

X coming out at least as high as the quantile $Q_X(p)$. (When a jump is present, the probability associated with the interval $[Q_X(p), \infty)$ may be greater than $1-p$, and F_X^p counters that by "spliting the probability atom" at $q = Q_X(p)$ appropriately,) For this reason, the terminology and notation

$$\overline{Q}_X(p) = \text{ conditional value-at-risk of } X \text{ at level } p \ = \mathrm{CVaR}_p(X) \tag{27}$$

was introduced in [11] as an extension of the terminology and notation

$$Q_X(p) = \text{ value-at-risk of } X \text{ at level } p \ = \mathrm{VaR}_p(X) \tag{28}$$

which is widely used for quantiles in finance when X represents potential loss in a portfolio of assets. Details covering possible jumps in F_X, such as the truncated distribution function F_X^p in (26), were provided in the follow-up paper [12]. (As already mentioned, "superquantile" was coined in [8] so that financial terminology would not be forced on users outside of finance.)

In modeling (22) by (23) in the case of $\mathcal{R}(X(u)) = \overline{Q}_{X(u)}(p)$ with $p = 0.95$, for example, the condition $f_{\mathcal{R}}(u) \leq b$ comes out as requiring that $X(u)$ be $\leq b$ "on average even in the worst 5% of cases." This is more conservative obviously than the quantile version, which would only require that $X(u)$ be $\leq b$ "except in the worst 5% of cases" and take no account of the seriousness of exceptional cases.

In contrast to the poor mathematical behaviour of $Q_{X(u)}(p)$ with respect to u, the superquantile $\overline{Q}_{X(u)}(p)$ is much nicer. An especially valuable property is the following: if the random variable $X(u)$ depends *convexly* on u, in the sense that

$$X((1-t)u + tu') \leq (1-t)X(u) + tX(u') \ \text{ almost surely for } t \in (0,1),$$

then $\overline{Q}_{X(u)}(p)$ is convex as a function of u. In other words, *the superquantile approach is convexity-preserving*, whereas the quantile approach might not even be continuity-preserving. In a vast range of applications in finance, at least, $X(u)$ depends not just convexly but *linearly* on u, and there superquantiles are all the more at home. Measures of risk based on superquantiles instead of quantiles also possess the property of *coherency* stressed in the pioneering work of Artzner et al. [1].

This good report about superquantiles could be unconvincing if the impression is held that the neither the "conditional expectation" description of $\overline{Q}_X(p)$ nor the integral formula equivalent to it in (21) is practical enough for implementation. However, there is yet another formula, first developed in [11] and extended in [12], which is eminently practical. For $p \in (0,1)$,

$$\overline{Q}_X(p) \ = \ \min_{C \in \boldsymbol{R}} \Big\{ C + \frac{1}{1-p} E[\, \max\{0, X - C\}\,] \Big\}, \tag{29}$$

whereas

$$[Q_X(p), Q_X^+(p)] \ = \ \operatorname*{argmin}_{C \in \boldsymbol{R}} \Big\{ C + \frac{1}{1-p} E[\, \max\{0, X - C\}\,] \Big\}. \tag{30}$$

In other words, quantile-superquantile pairs can be computed in tandem by solving a one-dimensional optimization problem involving a simplified expectation expression (with nothing "conditional").

This means moreover that, in the case of $X(u)$ depending on a decision vector u, the condition $\overline{Q}_{X(u)}(p) \leq b$ on u can be replaced by a condition on (u, C) where C is an additional decision variable, namely

$$C + \frac{1}{1-p} E[\, \max\{0, X - C\}\,] \leq b.$$

More about these features in optimization applications can be found in [11] and [12].

Returning now to the remarks at the end of the preceding section about conjugates of different choices of convex functions associated with the characteristic curve Γ_X, we point out that the minimization formula for $\bar{Q}_X(p)$ in (29) is obviously equivalent to the formula for the conjugate superexpectation function E_X^* in (20). In truth, this was the source of the discovery of that minimization formula in [11].

The general theory of measures of risk $\mathcal{R}$ is well developed by now with many contributions, but much still needs to be explored. A particularly interesting target could be the connections between the choice of $\mathcal{R}$ and the form of generalized regression that might be employed in treating the data on which the probabilistic underpinnings of an optimization model may depend. This line of research, initiated in [14], has been pushed further in the broad article [13] and on the other hand with more intense focus on superquantiles in [9] and [10].

Another topic connecting the convex analysis and duality associated with random variables X through their characteristic curves Γ_X is convergence of random variables. As proved in [9, Theorem 4], convergence *in distribution* of a sequence $\{X_k\}$ to X corresponds to *graphical* convergence of the corresponding characteristic curves Γ_{X_k} to Γ_X. Moreover it is equivalent to having $\overline{Q}_{X_k}(p) \to \overline{Q}_X(p)$ for every $p \in (0,1)$. Graphical convergence as an offshoot of set convergence is developed extensively in the book [15].

References

[1] Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D., "Coherent measures of risk," *Mathematical Finance* **9** (1999), 203–227.

[2] Lorenz, M., "Methods of measuring concentration of wealth," *Journal of the American Statistical Association* **9** (1905), 209–219.

[3] Minty, G.J., "Monotone networks," *Proceedings Royal Soc. London A* **257** (1960), 194–212.

[4] Ogryczak, W. and Ruszczyński, A., "Dual stochastic dominance and related mean-risk models," *SIAM Journal on Optimization* **13** (2002), 60–78.

[5] Rockafellar, R.T., *Convex Functions and Dual Extremum Problems*, doctoral dissertation, Harvard University, 1963.

[6] Rockafellar, R.T., *Convex Analysis*, Princeton University Press, 1970.

[7] Rockafellar, R.T., *Network Flows and Monotropic Optimization*, Wiley, 1983; republished by Athena Scientific, 1998.

[8] Rockafellar, R. T., and Royset, J. O., "On buffered failure probability in design and optimization of structures," *Journal of Reliability Engineering and System Safety* **99** (2010), 499–510.

[9] Rockafellar, R.T., and Royset, J.O., "Random variables, monotone relations and convex analysis," *Mathematical Programming B* [2014].

[10] Rockafellar, R.T., and Royset, J.O., "Superquantile regression with applications to buffered reliability, uncertainty quantification and conditional value-at-risk," European Journal of Operations Research **234** (2014), 140–154.

[11] Rockafellar, R. T. and Uryasev, S., "Optimization of conditional value-at-risk," *Journal of Risk* **2** (2000), 21–42.

[12] Rockafellar, R. T. and Uryasev, S., "Conditional value-at-risk for general loss distributions," *Journal of Banking and Finance* **26** (2002), 1443–1471.

[13] Rockafellar, R. T. and Uryasev, S. "The fundamental risk quadrangle in risk management, optimization and statistical estimation," *Surveys in Operations Research and Management Science* **18** (2013), 33–53.

[14] Rockafellar, R. T., Uryasev, S., and Zabarankin, M., "Risk tuning with generalized linear regression," *Math. of Operations Research* **33** (2008), 712–729.

[15] Rockafellar, R.T., and Wets, R. J-B, *Variational Analysis*, No. 317 in the series *Grundlehren der Mathematischen Wissenschaften*, Springer-Verlag, 1997.