AN ACTUARIAL INDEX OF THE RIGHT-TAIL RISK*

Shaun Wang†

ABSTRACT
A common characteristic for many insurance risks is the right-tail risk, representing low-frequency, large-loss events. In this paper I propose a measure of the right-tail risk by defining the right-tail deviation and the right-tail index. I explain how the right-tail deviation measures the right-tail risk and compare it to traditional measures such as standard deviation, the Gini mean, and the expected policyholder deficit. The right-tail index is applied to some common parametric families of loss distributions.

1. INTRODUCTION

In a broad sense, an insurance risk refers to the business, legal, or management aspects of transferring the economic impact of unforeseen mishaps. In this paper, the term “insurance risk” refers to a loss variable that quantifies the potential loss amount associated with an insurance contract or the whole book of an insurer’s business, depending upon the intended application. With this narrow definition, a characteristic of many insurance risks (individual or aggregate) is the right-tail risk, which represents low-frequency and large-loss events. This characteristic can be observed from the following two aspects:

1. Process Deviation. In liability insurance, insurance loss amounts or loss developments often are highly skewed and have long right tails. In property insurance, the hazards of natural disasters (earthquake, hurricane, flood) often manifest themselves at the right tail in a property writer’s aggregate claims distribution. In such situations, the large deviations due to the right-tail losses are a major concern to the insurer. Thus an indicator of the right-tail random deviation from the expected loss is desirable.

2. Parameter Risk. In practice, the probability distributions for losses or loss developments are seldom known with precision. There is always considerable uncertainty about the best-estimate probability distribution. In terms of statistical sampling error, the further at the right tail, the fewer are the available data and thus the higher the uncertainty regarding the best-estimate tail probabilities. For instance, in liability insurance, considerably greater uncertainty exists in increased-limits ratemaking than in the basic-limit ratemaking. Traditionally, the most commonly used measures of risks are variance and standard deviation. Standard deviation is a “standard” measure of deviation from the mean if the underlying variable has a normal distribution. Even though standard deviation has been used to measure the deviation from the mean for other than normal distributions, it is not a good risk measure for large insurance risks with skewed distributions. The poor performance of standard deviation in measuring insurance risks has been reported by many authors, for example, Ramsay (1993) and Lowe and Stanard (1996).

In clarification of the concept of risk margins, Philbrick (1994) discusses various types of risk measures as below. Here I elaborate Philbrick’s points with special emphasis on measures of the right-tail risk.

• “A risk margin based on a certainty equivalent concept.” The certainty equivalent is a well-established concept in actuarial and economic theories of risk and uncertainty (for example, the expected utility theory and premium calculation principles). The certainty equivalent may represent the market price for transferring the risk. Ideally, the certainty equivalent for a risk automatically adjusts for the right-tail risk. Therefore, once a certainty equivalent theory is established, a measure of the right-tail deviation can be derived from the difference between the certainty equivalent and the expected loss.

• “A risk margin based on probability confidence intervals.” This is a natural approach as long as the

---

*This research was supported by a grant from the Society of Actuaries Committee on Knowledge Extension Research.
†Shaun Wang, A.S.A., Ph.D., is an Associate Actuary at the SCOR Reinsurance Company, One Pierce Place, Itasca, Illinois 60143-4049, e-mail scorus/itasca/swang%5512559@mcimail.com.
parameter risk is of main concern. A measure of the right-tail risk should be capable of quantifying the parameter uncertainty.

- “A risk margin based on a theory of ruin.” The probability of ruin is essentially a quantile concept: (1) if a tolerable (small) probability of ruin is stated, it gives the amount of capital needed to ensure this safety level; (2) if a threshold amount of loss is stated, it gives the probability that the actual loss will exceed this threshold. Butsic (1994) proposes the use of the expected policyholder deficits (EPD) in determining risk-based-capital requirements. Gerber and Shiu (1997) discuss the probability of ruin as well as the deficit at ruin. The EPD is an advancement of the probability of ruin concept because it takes into account not only the probability of ruin but also the magnitude of deficit and thus reveals more information about the right tail than simply the probability of ruin. A good measure of the right-tail risk should utilize/reveal as much information as possible about the tail probabilities.

In this paper I propose a new measure of the right-tail deviation for a non-negative random variable X:

$$D[X] = \int_0^\infty \sqrt{\Pr(X > t)} \, dt - E[X].$$

I explain how \(D[X]\) measures the right-tail risk from various perspectives including (1) the certainty equivalent approach, (2) the parameter risk approach, and (3) distance between loss distributions. I show that \(D[X]\) preserves the basic properties of the standard deviation, namely,

- \(D[aX + b] = aD[X]\), for constants \(a > 0\) and \(b\).
- \(D[X + Y] \leq D[X] + D[Y]\) and equality holds for perfectly correlated risks.

However, the right-tail deviation differs from the standard deviation in the following aspects:

- The right-tail deviation \(D[X]\) preserves the common ordering of risks such as first and second stochastic dominance, while the standard deviation does not. Indeed, the right-tail deviation is much more capable of differentiating risks than standard deviation.
- The right-tail deviation \(D[X]\) is additive when a risk is divided into excess-of-loss layers, while the standard deviation is subadditive for layers. This distinctive behavior gives \(D[X]\) a comparative advantage in calculating risk charges in reinsurance pricing.

2. Preliminary Concepts

2.1 Expected Loss

For an insurance risk \(X\), a non-negative random variable, its cumulative distribution function is defined by \(F_X(t) = \Pr(X \leq t)\), and its decumulative distribution function (ddf) is defined by \(S_X(t) = \Pr(X > t)\). Whenever possible, we use the ddf representation, which has many advantages including a unified treatment of both continuous and discrete variables.

Lemma 2.1

For a non-negative random variable \(X\), we have

$$E[X] = \int_0^\infty S_X(t) \, dt.$$

This result can be found in Bowers et al. (1986), but here we give a simple yet general proof.

Proof

For \(x \geq 0\) it is true that

$$x = \int_0^\infty I(x > t) \, dt,$$

where \(I\) is the indicator function. For a non-negative random variable, it holds that

$$X = \int_0^\infty I(X > t) \, dt.$$

By taking expectation on both sides of the equation, we get

$$E[X] = \int_0^\infty E[I(X > t)] \, dt = \int_0^\infty S_X(t) \, dt. \quad \square$$

2.2 Insurance Layers

Most insurance contracts have some policy provisions such as deductibles and limits. A large risk is often divided into layers or quota shares among several insurers. Here we introduce a general term of layers. A layer \((a, a + h]\) of risk \(X\) is defined as an excess-of-loss cover:

$$X_{[a,a+h]} = \begin{cases} 0, & 0 \leq X < a \\ (X - a), & a \leq X < a + h, \\ h, & a + h \leq X \end{cases}$$

where \(a\) is the deductible (or retention), and \(h\) is the limit.

The following result can be easily verified.
Lemma 2.2

The excess-of-loss cover $X_{[a, a+h]}$ has a ddf:

$$S_{X_{[a, a+h]}}(t) = \begin{cases} S_X(a + t), & t < h \\ 0, & t \geq h. \end{cases}$$

From Lemmas 2.1 and 2.2, the expected loss for the layer $(a, a+h)$ is

$$E[X_{[a, a+h]}] = \int_0^a S_{X_{[a, a+h]}}(t) \, dt = \int_0^h S_X(a + u) \, du = \int_a^{a+h} S_X(t) \, dt.$$

Remark

A layer $(a, a+h)$ is like a window. The payment of a layer $(a, a+h)$ is the part of insurance claims that is observed through this window. Also, the ddf for a layer $(a, a+h)$ is the part of the original ddf that falls into this window.

Remark

Note that $S_X(t)$ represents the frequency of hitting the layer $(t, t+dt]$, and $S_{X_{[a, a+h]}}(t) \, dt$ represents the expected loss for the infinitesimal layer $(t, t+dt]$. Furthermore, $X_{(t,t\downarrow dt]}$ has approximately a Bernoulli distribution with

$$\Pr\{X_{(t,t\downarrow dt]} = 0\} = 1 - S_X(t), \quad \Pr\{X_{(t,t\downarrow dt]} = dt\} = S_X(t).$$

2.3 Comonotonicity

Traditionally, the relationship of two random variables, $X$ and $Y$, is usually measured by the Pearson correlation coefficient:

$$\rho(X, Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}},$$

where $\text{Cov}(X,Y) = E[XY] - E[X]E[Y]$ is the covariance of $X$ and $Y$. If $\rho(X, Y) = 0$, then $X$ and $Y$ are uncorrelated. If $\rho(X, Y) = 1$, then $X$ and $Y$ are perfectly correlated; in this case it is necessary that $Y = aX + b$ with $a > 0$.

The concept of perfect correlation is too restrictive because it holds only for random variables with linear relationships $Y = aX + b$ with $a > 0$. As a generalization of perfect correlation, an important concept of comonotonicity introduced by Yaari (1987) and Schmeidler (1986) has played a very important role in decision theory under uncertainty.

Definition 2.1

Two risks $X$ and $Y$ are comonotonic if there exists a random variable $Z$ and nondecreasing real functions $u$ and $v$ such that

$$X = u(Z), \quad Y = v(Z), \quad \text{with probability one.}$$

Comonotonicity is a generalization of the concept of perfect correlation to random variables without linear relationships. Note that perfectly correlated risks are comonotonic, but its converse does not hold. Consider two layers $(a, a+h)$ and $(b, b+h)$ for a continuous variate $X$. The layer payments $X_{[a, a+h]}$ and $X_{[b, b+h]}$ are comonotonic because both are nondecreasing functions of the original risk $X$. They are bets on the same event, and neither of them is a hedge against the other. On the other hand, for $a \neq b$, $X_{[a, a+h]}$ and $X_{[b, b+h]}$ are not perfectly correlated because neither can be written as a linear function of the other.

Lemma 2.3

For two comonotonic risks $X$ and $Y$, $\text{Cov}(X,Y) \geq 0$.

Proof

If $X$ and $Y$ are comonotonic, from Definition 2.1 there exists a random variable $Z$ such that

$$X = u(Z), \quad Y = v(Z), \quad \text{with probability one,}$$

where the functions $u$ and $v$ are nondecreasing. Since $u$ is nondecreasing, there exists a number $t_0$ such that $u(t) \geq E[u(Z)]$ for $t > t_0$ and $u(t) \leq E[u(Z)]$ for $t < t_0$. Now we have

$$E[XY] - E[X]E[Y] = \int_0^a u(t)v(t) \, dF_Z(t)$$

\begin{align*}
&= \int_0^{t_0} \{u(t) - E[u(Z)]\} \, v(t) \, dF_Z(t) + \int_{t_0}^a \{u(t) - E[u(Z)]\} \, v(t) \, dF_Z(t) \\
&\geq \int_0^{t_0} \{u(t) - E[u(Z)]\} \, v(t_0) \, dF_Z(t) + \int_{t_0}^a \{u(t) - E[u(Z)]\} \, v(t_0) \, dF_Z(t) \\
&\geq 0. \quad \Box
\end{align*}
2.4 Additivity

For a risk measure it is often desirable to have some sort of additivity criterion for aggregating/allocating the total risk measure of a risky portfolio.

In many situations, a natural requirement for a risk measure is subadditivity, that is,

\[ R[X_1 + X_2] \leq R[X_1] + R[X_2]. \]

Artzner et al. (1996) argue for this subadditivity in the context of value at risk or capital requirement:

If a risk measure, for example, were to fail to satisfy this property, then an individual would be motivated to open two accounts, one holding \( X \), and the other \( Y \), incurring the smaller risk measure of \( R[X_1] + R[X_2] \). Similarly, if a firm were required to meet a capital requirement which did not satisfy this property, it would be motivated to set up two subsidiaries.

While subadditivity may reflect risk reduction through risk pooling, additivity may be required if the risk pooling effect does not exist. Because there is no hedge among comonotonic risks, it would be desirable to require that a risk measure be additive for comonotonic risks. Recall that insurance layers of the same risk are comonotonic; thus it is desirable for risk measure to be additive for insurance layers.

2.5 Ordering of Risks

There are widely accepted methods for comparing risks in the economic and statistical theories of risk. The most basic concepts are the first and second stochastic dominance.

A risk \( X \) is smaller than risk \( Y \) in the first stochastic dominance (notation \( X <_{1st} Y \)) if any of the following equivalent conditions is met (Hadar and Russell 1969):

1. For every decision-maker who has an increasing concave function \( u \):
   \[ E[u(-X)] \geq E[u(-Y)], \]
   that is, a common ordering shared by all risk-averse individuals.
2. \[ \int_x^\infty S_x(t) \, dt \leq \int_x^\infty S_y(t) \, dt, \text{ for all } x \geq 0, \]
   that is, the net expected loss for the layer \( X_t \) is always higher with risk \( Y \). For this reason, second stochastic dominance is also called stop-loss ordering in the actuarial literature (Kaas et al. 1994).
3. \( Y \geq X + Z \) in which \( E[Z|X] \geq 0 \) with probability one, that is, \( Y \) is equal (in distribution) to \( X \) plus noise \( Z \).

Now we apply these concepts in comparing the relative riskiness of different layers of the same risk.

Lemma 2.4

Given a risk \( Y \),

- For two layers \( (a, a+h) \) and \( (b, b+h) \) with the same limit \( h \),
  \[ a < b \Rightarrow X_{(b,b+h)} <_{1st} X_{(a,a+h)} \]
- For two layers \( (a, a+h) \) and \( (b, b+h) \) with the same expected loss, that is, \( E[X_{(a,a+h)}] = E[X_{(b,b+h)}] \),
  we have
  \[ X_{(a,a+h)} <_{2nd} X_{(b,b+h)} \]

Proof

See Wang (1996a, pp. 74–75).

Now we are ready to move on the discussions of a right-tail risk measure.

3. The Certainty Equivalent Approach

For a risk \( X \), let

\[ H[X] = E[X] + R[X] \]

represent the certainty equivalent to risk \( X \), or the price for transferring the risk \( X \) to other parties. We are hoping that a theory on the certainty equivalent \( H[X] \) may induce a measure of the right-tail risk, \( R[X] \).

Wang, Young, and Panjer (1997) discuss the axiomatic characterization of insurance prices. They
propose five basic axioms for insurance prices and give interpretations for each of them.

Axiom 1. For a given market condition, the price of an insurance risk $X$ depends only on its distribution. That is, if $S_x = S_y$, then $H[X] = H[Y]$.

Axiom 2. If $X \leq Y$ with probability one, then $H[X] \leq H[Y]$.

Axiom 3. If $X$ and $Y$ are comonotonic, then $H[X + Y] = H[X] + H[Y]$.

Axiom 4. For $d \geq 0$, $\lim_{d \to 0^+} H[X(d,1)] = H[X]$, and $\lim_{d \to \infty} H[\min(X,d)] = H[X]$.

Theorem 3.1 (Wang, Young, and Panjer 1997)

If the certainty equivalent $H[X]$ satisfies Axioms 1–4 and $H[1] = 1$, then $H$ has the following representation:

$$H[X] = \int_0^\infty g[S_x(t)] \, dt,$$

where $g$ is a distortion function (that is, increasing with $g(0) = 0$ and $g(1) = 1$).

Furthermore, the $H[X]$ in Theorem 3.1 has the following properties

- $H[aX + b] = aH[X] + b$ for $a, b \geq 0$.
- $H[X] \geq E[X]$ if and only if $g(x) \geq x$ for all $x \in [0,1]$.
- $H$ preserves the second stochastic dominance if and only if $g$ is concave.
- If $g$ is concave, then $H[X + Y] \leq H[X] + H[Y]$.

As a by product of this certainty equivalent approach, we obtain a general class of risk measures for the right-tail deviation:

$$R[X] = \int_0^\infty g[S_x(t)] \, dt - E[X],$$

where $g$ is increasing concave with $g(0) = 0$ and $g(1) = 1$.

In addition to the four basic axioms for insurance prices, Wang, Young, and Panjer (1997) also propose the following axiom for reduction of compound Bernoulli risks and give a no-arbitrage interpretation.

Axiom 5. Let $Y = BX$ be a compound Bernoulli risk, where the Bernoulli frequency random variable $B$ is independent of the loss severity random variable $X = Y|Y > 0$. Then the market prices for risks $Y = BX$ and $BH[X]$ are equal.

Theorem 3.2 (Wang, Young, and Panjer 1997)

If the insurance price $H[X]$ satisfies Axioms 1–5, then we have the following unique representation:

$$H[X] = \int_0^\infty [S_x(t)]^r \, dt,$$

where $0 \leq r$. Furthermore, if $H[X] \geq E[X]$, then $0 \leq r \leq 1$.

Besides Theorem 3.2, there are many other reasons to use distortions $g$ of the form $g(x) = x^r$, $0 \leq r \leq 1$; see Wang (1996a, b) for further details. In order to yield a sharp numerical indicator for the right-tail risk, in this paper we choose the square root function: $g(x) = \sqrt{x}$.

4. The Right-Tail Deviation

As a consequence of the certainty equivalent theory in Theorems 3.1 and 3.2, we introduce a new risk measure for the right-tail deviation.

Definition 4.1

For a non-negative random variable $X$ with decumulative distribution function $(ddf) S_x(t) = \Pr(X > t)$, we define the right-tail deviation as

$$D[X] = \int_0^\infty \sqrt{S_x(t)} \, dt - \int_0^\infty S_x(t) \, dt,$$

and a right-tail index as

$$d(X) = \frac{D[X]}{E[X]} = \frac{\int_0^\infty \sqrt{S_x(t)} \, dt}{\int_0^\infty S_x(t) \, dt} - 1.$$

It is straightforward to show that the right-tail deviation $D[X]$ satisfies:

- If $\Pr(X = b) = 1$, then $D[X] = 0$.
- Scale invariant: $D[cX] = cD[X]$ for $c > 0$.
- Shift invariant: $D[X + b] = D[X]$ for any constant $b$.
- If $X$ and $Y$ are comonotonic (including perfect correlation), then $D[X + Y] = D[X] + D[Y]$.

Proposition 4.1

For a small layer $(t, t + dt)$, we have

1. $D[X_{t,t+dt}] \leq \sigma[X_{t,t+dt}]$
2. For a small layer $(t, t + dt)$, the ratio

$$\frac{D[X_{t,t+dt}]}{\sigma[X_{t,t+dt}]}$$

is an increasing function of $t$.
3. If $X$ is a continuous variable with support over $[0, \infty)$ (or can be infinity),

$$\lim_{t \to \infty} \frac{D[X_{t,t+dt}]}{\sigma[X_{t,t+dt}]} = 1.$$
4. For a non-negative random variable $X$, the right-tail deviation $D[X]$ is finite if, and only if, the standard deviation $\sigma(X)$ is finite.

Proof
Let $u = S_x(t)$ be the probability of hitting the layer $[t, t+dt]$. The layer payment from $[t, t+dt]$ has approximately a Bernoulli distribution:

$$\Pr\{X_{[t,t+dt]} = 0\} = 1 - u, \quad \Pr\{X_{[t,t+dt]} = dt\} = u.$$ 

Thus,

$$\sigma[X_{[t,t+dt]}] = \sqrt{u - u^2} \, dt.$$ 

On the other hand,

$$D[X_{[t,t+dt]}] = (\sqrt{u} - u) \, dt.$$ 

We can show that (1) $\sqrt{u} - u \leq \sqrt{u - u^2}$ for $0 \leq u \leq 1$, (2) the ratio $(\sqrt{u} - u)/(\sqrt{u - u^2})$ increases as $u$ decreases, and (3) $\lim_{u \to 0} \frac{\sqrt{u} - u}{\sqrt{u - u^2}} = 1$. 

(4) This is based on a formula in Aebi/Embrechts/Mikosch (1992, pp. 147, Remarks (ii)).

5. **Gini Index**

Historically, some long-tailed distributions have an origin in income distributions, for example, Pareto and lognormal distributions; see Arnold (1983). In social welfare studies, a celebrated measure for income inequality is the Gini index.

**Definition 5.1**
Assume that level of wealth for all individuals in a country (community) can be summarized by a distribution, $S_x(u)$, representing the proportion of individuals with wealth in excess of $u$. As a measure of income inequality of a society, the Gini index is defined as

$$\text{gini}(X) = \frac{\mathbb{E}[(X_1 - X_2)^2]}{2\mathbb{E}[X]},$$

where $X_1$ and $X_2$ are independent and have the same distribution as $X$.

**Proposition 5.1**
The Gini index can be equivalently represented as

$$\text{gini}(X) = 1 - \frac{\int_0^\infty [S_x(u)^2] \, du}{\int_0^\infty S_x(u) \, du}.$$

The following proof is due to Dorfman (1979).

In other words, for a small layer at the right tail, the standard deviation and the right-tail deviation converge to each other, as demonstrated in the following example.

**Example 4.1**
Consider the claim distribution with a ddf:

$$S_x(t) = \left(\frac{1000}{1000 + t}\right)^2.$$ 

For different layers $[a, a+h]$ with fixed limit $h=1000$, we compare the standard deviation and the right-tail deviation in Table 1.

<table>
<thead>
<tr>
<th>Layer, $L$</th>
<th>Expected Loss, $E[L]$</th>
<th>Standard Deviation, $\sigma(L)$</th>
<th>Right-tail Deviation, $D[L]$</th>
<th>Percentage Difference, $\sigma(L)/D[L] - 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(0, 1000]$</td>
<td>500.0</td>
<td>369.2</td>
<td>193.1</td>
<td>91.1%</td>
</tr>
<tr>
<td>$(1000, 2000]$</td>
<td>166.7</td>
<td>341.3</td>
<td>238.8</td>
<td>42.9</td>
</tr>
<tr>
<td>$(10000, 11000]$</td>
<td>7.576</td>
<td>85.43</td>
<td>79.44</td>
<td>7.55</td>
</tr>
<tr>
<td>$(100000, 101000]$</td>
<td>0.99707</td>
<td>9.836</td>
<td>9.755</td>
<td>0.83</td>
</tr>
<tr>
<td>$(1000000, 1001000]$</td>
<td>0.0009970</td>
<td>0.9983</td>
<td>0.9975</td>
<td>0.08</td>
</tr>
<tr>
<td>$(10000000, 10001000]$</td>
<td>$9.9997 \times 10^{-6}$</td>
<td>0.09998</td>
<td>0.09998</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Here "income inequality" refers to the polarization of the wealth distribution.
role of the right-tail index in their definition formula. Later we show that the right-tail risk is parallel to the role of the Gini index in measuring income inequalities.

6. Probability Metrics

In the study of various limit theorems in probability theory, probability metrics have been developed to measure the distance between probability distributions; see Zolotarev (1979) and Rachev (1991). Here we introduce one of the most popular probability metrics, namely, the Kantorovich metric as given by Rachev (1991, pp. 27-28), or Mallows metric as given in the actuarial literature by Aebi, Embrechts and Mikosch (1992, pp. 144-145).

Definition 6.1

For two decumulative distribution functions $S_1$ and $S_2$ on $\mathbb{R}$, the first order (Kantorovich-) Mallows metric is defined as

$$M(S_1, S_2) = \inf \{E[|X - Y|] : S_X = S_1, \quad S_Y = S_2 \}.$$ 

Aebi, Embrechts and Mikosch (1992, pp. 144-145) indicate that the notion of Mallows metric has many applications in insurance mathematics. They show the following results:

Proposition 6.1

Let $U$ be a uniformly distributed random variable on $[0,1]$. For two decumulative distribution functions $S_1$ and $S_2$, the first-order Mallows metric can be calculated as:

$$M(S_1, S_2) = E[|S^{-1}_1(U) - S^{-1}_2(U)|]$$

$$= \int_0^1 |S_1(t) - S_2(t)| dt.$$ 

In other words, for any two distributions, their Mallows metric is the expected absolute difference of two comonotonic variables with the respective distributions. [For a detailed discussion on comonotonicity, see Wang and Dhaene (1997).]

For a non-negative random variable $X$, we can easily see that the right-tail deviation has a simple representation in terms of Mallows metric:

$$D[X] = M(S_X, S_X^R),$$

and so does the Gini mean:

$$G[X] = M(S_X, S_X^\theta).$$
7. TRADITIONAL METHODS

In the actuarial literature there have been many discussions on risk measures in the context of premium calculation principles (for example, Goovaerts et al. 1984). However, most traditional methods do not satisfy layer additivity. In this section, we analyze how those risk measures behave when layering a risk. Interestingly, by forcing them to be additive for layers, we recover the right-tail deviation and Gini mean.

7.1 Variance and Standard Deviation

Variance and standard deviation are the most widely used risk measures (Bowers et al. 1986). One criticism of the variance or standard-deviation based risk measures is that they do not preserve the first stochastic dominance; see Kaas et al. (1994, p. 17).

Here we give a detailed analysis of how they behave when dividing a risk into layers. Neither variance nor measures is that they do not preserve the first stochastic dominance. For any non-negative random variable $X$ with $X \geq 0$ and support over $[0, \infty)$, we have in general that
\[
\text{Var}(X) = \text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2) + 2 \text{Cov}(X_1, X_2).
\]
From Lemma 2.3, we have $0 \leq \text{Cov}(X_1, X_2) \leq \sigma(X_1) \sigma(X_2)$, and thus
\[
\text{Var}(X) \geq \text{Var}(X_1) + \text{Var}(X_2)
\]
but
\[
\sigma(X) \leq \sigma(X_1) + \sigma(X_2).
\]

Consider a risk $X$ with a finite second moment. A risk $X \geq 0$ can be divided into many small layers
\[
X = \sum_{j=0}^{\infty} X_{j \leftarrow j+1/h}, \quad h > 0.
\]
As the division refines (that is, $h \to 0$), the small layer $X_{j \leftarrow j+1/h}$ has approximately a Bernoulli distribution:
\[
\Pr \{X_{j \leftarrow j+1/h} = 0 \} \approx 1 - S_X(jh),
\]
\[
\Pr \{X_{j \leftarrow j+1/h} = h \} \approx S_X(jh).
\]
Thus,
\[
\text{Var}(X_{j \leftarrow j+1/h}) \approx h^2 S_X(jh)(1 - S_X(jh)).
\]
Therefore, the sum of the variances of layers approaches zero, indeed,
\[
\sum_{j=0}^{\infty} \text{Var}(X_{j \leftarrow j+1/h}) \sim h \int_0^\infty \left[ S_X(t) - S_X(t)^2 \right] dt = hG(X).
\]

On the other hand, the standard deviation exhibits subadditivity. By dividing a risk into pieces of smaller risks, the total standard deviation would generally increase:
\[
\sigma(X_1 + X_2) \leq \sigma(X_1) + \sigma(X_2).
\]
It is easy to verify that a small layer $X_{j \leftarrow j+1/h}$ has a standard deviation
\[
\sigma(X_{j \leftarrow j+1/h}) \approx h \sqrt{S_X(jh)[1 - S_X(jh)]}.
\]
Therefore, by dividing a risk $X$ into many small layers $X_{j \leftarrow j+1/h}$ and taking the limit as the length of the layer $h$ goes to zero, the sum of the standard deviations of layer payments approaches to a maximum
\[
\sum_{j=0}^{\infty} \sigma(X_{j \leftarrow j+1/h}) \to \int_0^\infty \sqrt{S_X(t)[1 - S_X(t)]} dt.
\]

Definition 7.1

For any non-negative random variable $X$ with $S_X(t) = \Pr(X \geq t)$, we define the maximal standard deviation as
\[
\text{MSD}[X] = \int_0^\infty \sqrt{S_X(t)[1 - S_X(t)]} dt.
\]

The following results can be easily verified:

- The maximal standard deviation is additive for layers.
- $D[X] \leq \text{MSD}[X]$.
- If $X$ is a continuous variable with support over $[0, \omega)$ ($\omega$ can be infinity),
\[
\lim_{t \to \omega} \frac{D[X_{t \leftarrow t}]}{\text{MSD}[X_{t \leftarrow t}]} = 1.
\]

- Unlike the right-tail deviation, the maximal standard deviation does not preserve first or second stochastic dominance.

The same extent, for the proponents of the standard deviation, the right-tail deviation can be viewed as a simplified or improved version of the maximum standard deviation.

7.2 Quantile Concept

Sometimes management is concerned only with some threshold amount of loss, and quantiles (say, 95th or 99th percentile) can serve as a simple risk indicator.
It does have some drawbacks because it fails to reveal more information at other percentiles. Some authors point out that subadditivity does not always hold for quantiles; see Artzner et al. (1996) and Gerchak and Mossman (1992). As an advancement from the quantile concept, Butsic (1994) proposes the use of expected policyholder deficit in deciding risk-based-capital requirements.

7.3 Expected Policyholder Deficit

Let the threshold amount of loss be $\beta E[X]$, $\beta > 0$. The expected policyholder deficit (EPD) in excess of $\beta E[X]$ is defined as

$$\text{EPD}_\beta[X] = E[X_{(\beta E[X])}] = \int_{\beta E[X]}^{\infty} S_X(t) \, dt.$$  

Van Heerwaarden and Kaas (1992) propose a Dutch premium calculation principle given by

$$H[X] = E[X] + \alpha \text{EPD}_\beta[X], \quad 0 \leq \alpha \leq 1.$$  

They show that $H[X]$ is subadditive and preserves the second stochastic dominance. However, Wang (1996c, p. 111) gives an example that shows that EPD(\(\beta\)) is not additive for layers.

Now we divide a risk $X \geq 0$ into many small layers, $X_{(j+1)h} \leq t < X_{jh}$, $h > 0$, and consider the EPD with $\beta = 1$. Note that for the small layer $(t, t+dt)$ we have

$$E[X_{(tt+dt)}] = S_X(t) \, dt,$$
$$\text{EPD}_1[X_{(tt+dt)}] = [1 - S_X(t)]S_X(t) \, dt.$$  

Therefore, by dividing a risk $X \geq 0$ into many small layers and taking a limit as the length of the layer goes to zero, we get a maximal (total) EPD at

$$\int_0^\infty [1 - S_X(t)]S_X(t) \, dt = G[X].$$  

7.4 The $p$-th Mean Value

Goovaerts et al. (1984) discuss the method of calculating insurance premiums using the $p$-th mean value

$$E_p[X] = \left( E[X^p] \right)^{1/p}, \quad p \geq 1.$$  

They also show that $E_p[X]$ is subadditive.

By dividing a risk $X \geq 0$ into many small layers, $X_{(j+1)h} \leq t < X_{jh}$, we can show that the total $p$-th mean value approaches a maximum

$$\sum_{j=0}^\infty E_p[X_{(j+1)h}] \to \int_0^\infty [S_X(t)]^{1/p} \, dt,$$  

which gives the PH-transform principle of Wang (1995). This induces a measure of deviation from the mean:

$$R_p[X] = \int_0^\infty [S_X(t)]^{1/p} \, dt - E[X],$$  

where $p=2$ corresponds to our right-tail deviation.

8. Examples

Now we give some examples of how to calculate the right-tail index and the Gini index.

Example 8.1

For a Bernoulli risk with $\Pr\{X=0\} = 1-q$ and $\Pr\{X=1\} = q$, the right-tail deviation is $D[X] = q-1$ and the Gini mean is $G[X] = q-(q-1)$. Accordingly, the right-tail index is $d(X) = 1/\sqrt{q-1}$ and the Gini index $gini(X) = 1-q$.

Similarly, for an arbitrary risk $X$, the small layer $X_{(tt+dt)}$ can be approximated as a Bernoulli variable. Thus,

$$d(X_{(tt+dt)}) = 1/\sqrt{S_X(t)} - 1$$
and
$$gini(X_{(tt+dt)}) = 1 - S_X(t).$$

Example 8.2

Assume that $X$ has an exponential distribution with mean $\lambda$, that is, with ddf:

$$S_X(t) = e^{-t/\lambda}.$$  

(a) It can be easily verified that both the right-tail index and the coefficient of variation are equal to 1 and the Gini index is 1/2.

(b) For a layer $(0, x]$ of $X$, the right-tail index is

$$d(X_{(0,x]}) = [1 - e^{-x/\lambda}]^2.$$  

Example 8.3

Assume that $X$ is Pareto distributed with ddf:

$$S_X(t) = \left( \frac{\lambda}{\lambda + t} \right)^a, \quad t \geq 0.$$  

The right-tail index and the Gini index are (respectively):

\[\text{Example 8.4}\]

Assume that $X$ is lognormal distributed with

$$S_X(t) = \exp\left( \mu t - \frac{\sigma^2 t^2}{2} \right)$$  

The right-tail index and the Gini index are (respectively):

\[\text{Example 8.5}\]

Assume that $X$ is gamma distributed with

$$S_X(t) = \left( 1 + \frac{t}{\alpha} \right)^{-\lambda}$$  

The right-tail index and the Gini index are (respectively):
\[ d(X) = \begin{cases} \frac{\alpha}{(\alpha - 1)(\alpha - 2)}, & \alpha > 2 \\ \infty, & \alpha \leq 2. \end{cases} \]

\[ \text{gini}(X) = \begin{cases} \frac{\alpha}{2\alpha - 1}, & \alpha > 1 \\ 1, & \alpha \leq 1. \end{cases} \]

Example 8.4

Consider the exponential-inverse gaussian (E-IG) distribution with ddf:

\[ S(x) = \exp\left\{ \frac{\mu}{\beta} [1 - (1 + 2\beta x)^{1/2}] \right\}, \quad x > 0. \]

The E-IG distribution has the following mean and standard deviation (Hesselager, Wang, and Willmot 1997):

\[ E[X] = \frac{\beta + \mu}{\mu^2}, \quad \sigma(X) = \frac{\sqrt{5\beta^2 + 4\mu + \mu^2}}{\mu^2}. \]

It can be easily verified that

\[ D[X] = \frac{3\beta + \mu}{\mu^2}, \quad d(X) = 1 + \frac{2\beta}{\beta + \mu}. \]

Note that the right-tail deviation is greater than the standard deviation. The Gini index is

\[ \text{gini}(X) = \frac{1}{2} + \frac{\beta}{4(\beta + \mu)}. \]

Example 8.5

As we have seen in Example 8.2, for an exponential distribution, the right-tail index and the coefficient of variation are all equal to 1, while the Gini index is one half. Now we want to investigate how the right-tail index behaves for probability distributions with thicker or thinner tails than an exponential distribution. To do so, we investigate the gamma(\( \alpha \)) distribution with pdf

\[ f(x) = \frac{\lambda^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}, \quad x \geq 0, \]

which has a mean of \( \alpha \) and a coefficient of variation at \( 1/\sqrt{\alpha} \).

- When \( \alpha > 1 \), it has a thinner tail than an exponential distribution (asymptotically).
- When \( \alpha = 1 \), it is an exponential distribution.
- When \( \alpha < 1 \), it has a thicker tail than an exponential distribution.

In Table 2, we compute the coefficient of variation, the right-tail index, and the Gini index for different values of \( \alpha \).

<table>
<thead>
<tr>
<th>Gamma (( \alpha )) Distribution</th>
<th>Coefficient of Variation</th>
<th>Right-Tail Index</th>
<th>Gini Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha = 5 )</td>
<td>0.447</td>
<td>0.381</td>
<td>0.246</td>
</tr>
<tr>
<td>( \alpha = 3 )</td>
<td>0.577</td>
<td>0.514</td>
<td>0.313</td>
</tr>
<tr>
<td>( \alpha = 2 )</td>
<td>0.707</td>
<td>0.656</td>
<td>0.375</td>
</tr>
<tr>
<td>( \alpha = 1 )</td>
<td>1.000</td>
<td>1.000</td>
<td>0.500</td>
</tr>
<tr>
<td>( \alpha = 1/2 )</td>
<td>1.414</td>
<td>1.532</td>
<td>0.637</td>
</tr>
<tr>
<td>( \alpha = 1/3 )</td>
<td>1.732</td>
<td>1.963</td>
<td>0.713</td>
</tr>
<tr>
<td>( \alpha = 1/5 )</td>
<td>2.236</td>
<td>2.671</td>
<td>0.798</td>
</tr>
</tbody>
</table>

We can see that, as the right-tail becomes thicker (thinner), our right-tail index \( d(X) \) increases (decreases) faster than the coefficient of variation, which in turn is faster than the Gini index.

Note that the sum of \( k \) independent gamma(1), that is, exponential, variables has a gamma(\( k \)) distribution. From Table 2 we observe that the pooling of independent exponential variables results in more reduction in the right-tail index than the coefficient of variation.

8.1 Comparison of Continuous Distributions

Consider the following loss distributions each with the same mean (= 1) and variance (= 3).

- A Pareto distribution with \( \lambda = 2 \) and \( \alpha = 3 \) (see Example 8.3).
- An E-IG distribution with \( \mu = (3 + \sqrt{5})/2 \) and \( \beta = 2 + \sqrt{5} \) (see Example 8.4).
- A lognormal distribution with ddf

\[ S(t) = \Psi(\ln t - \mu/\sigma), \quad t > 0, \]

where \( \mu = -\ln 2, \quad \sigma = \sqrt{\ln 4} \) and \( \Psi(.) \) is the ddf for a standard normal distribution.

- A gamma distribution with pdf

\[ f(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}, \quad t > 0, \]

where \( \alpha = \lambda = 1/3 \).

- An inverse-Gaussian distribution with pdf

\[ f(t) = \mu(2\pi\beta^2)^{-1/2} \exp\left\{ -\frac{(t - \mu)^2}{2\beta^2} \right\}, \quad t > 0, \]

where \( \mu = 1 \) and \( \beta = 3 \).
A Weibull distribution with a df $S(t) = \exp \{-\alpha t^\beta\}$, where $\alpha=0.607248$ and $\beta=1.26957$.

Without referring to higher moments, we can order them by the right-tail index $d(X)$.

**Table 3**

<table>
<thead>
<tr>
<th>Continuous Distributions</th>
<th>Mean</th>
<th>Coefficient of Variation</th>
<th>Right-Tail Index</th>
<th>Gini Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pareto</td>
<td>1</td>
<td>$\sqrt{3}$</td>
<td>3.00</td>
<td>0.600</td>
</tr>
<tr>
<td>Lognormal</td>
<td>1</td>
<td>$\sqrt{3}$</td>
<td>2.59</td>
<td>0.595</td>
</tr>
<tr>
<td>E-IG</td>
<td>1</td>
<td>$\sqrt{3}$</td>
<td>2.24</td>
<td>0.655</td>
</tr>
<tr>
<td>Inverse-Gaussian</td>
<td>1</td>
<td>$\sqrt{3}$</td>
<td>2.17</td>
<td>0.632</td>
</tr>
<tr>
<td>Weibull</td>
<td>1</td>
<td>$\sqrt{3}$</td>
<td>2.13</td>
<td>0.681</td>
</tr>
<tr>
<td>Gamma</td>
<td>1</td>
<td>$\sqrt{3}$</td>
<td>1.96</td>
<td>0.713</td>
</tr>
</tbody>
</table>

For these parametric distributions, the ranking by the right-tail index is in agreement with our knowledge of their relative tail thickness; see Embrechts and Veraverbeke (1982) and Panjer and Willmot (1992, p. 350). However, the Gini indices are not in agreement with the commonly perceived ranking of tail thickness.

In summary, as its name may suggest, the right-tail deviation measures the right-tail risk, as opposed to the standard deviation, which measures the deviation about the mean, and as opposed to the Gini index, which measures the polarization of the wealth distribution.

**8.2 Comparison of Counting Distributions**

In the same way that the right-tail index $d(X)$ is used for claim severity distributions, we can compare claim frequency distributions. Here we consider the most popular two-parameter counting distributions, namely:

- A two-point Bernoulli distribution with a probability function $p(0)=3/4$, and $p(1)=1/4$.
- Negative binomial distribution with a probability function:

$$p_n = \frac{\Gamma(r+n)}{\Gamma(r) \, n!} \left( \frac{1}{1+\beta} \right)^r \left( \frac{\beta}{1+\beta} \right)^n, \quad n = 0, 1, 2, \ldots$$

where $r=0.5$ and $\beta=2$.
- Poisson inverse gaussian distribution with a probability generating function:

$$P(z) = \sum_{n=0}^{\infty} p_n z^n = \exp \left\{ \frac{\mu}{\beta} - \frac{\mu}{\beta} \sqrt{1 - 2\beta (z-1)} \right\},$$

where $\mu=1.0$ and $\beta=2.0$. Note that the probabilities can be evaluated recursively, see Willmot (1987, pp. 114–115).

- Generalized Poisson distribution with a probability function (see Consul 1990):

$$p_n = \theta (\theta + n \lambda)^{n-1} \frac{e^{-\theta n \lambda}}{n!}, \quad n = 0, 1, 2, \ldots$$

where $\theta=1/\sqrt{3}$ and $\lambda=1-1/\sqrt{3}$.

The parameters were chosen so that all three distributions have the same mean ($=1$) and variance ($=3$). Their right-tail indices and Gini indices can be easily computed, as listed in Table 4.

**Table 4**

<table>
<thead>
<tr>
<th>Discrete Distributions</th>
<th>Mean</th>
<th>Coefficient of Variation</th>
<th>Right-Tail Index</th>
<th>Gini Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poisson Inverse Gaussian</td>
<td>1</td>
<td>$\sqrt{3}$</td>
<td>2.01</td>
<td>0.72</td>
</tr>
<tr>
<td>Generalized Poisson</td>
<td>1</td>
<td>$\sqrt{3}$</td>
<td>1.94</td>
<td>0.73</td>
</tr>
<tr>
<td>Negative Binomial</td>
<td>1</td>
<td>$\sqrt{3}$</td>
<td>1.87</td>
<td>0.74</td>
</tr>
<tr>
<td>Bernoulli</td>
<td>1</td>
<td>$\sqrt{3}$</td>
<td>1.00</td>
<td>0.75</td>
</tr>
</tbody>
</table>

The right-tail index ranks P-IG as having a fatter tail than generalized Poisson, which in turn has a fatter tail than the negative binomial distribution. This ranking is consistent with their asymptotic behavior; see Willmot (1997) and Consul (1990). However, the Gini index indicates a reversed ordering. This suggests that the Gini index is not a good measure for the right-tail risk.

**9. Applications in Risk Charges**

Similar to the standard deviation principle, we can use the right-tail deviation in deciding risk-charges:

$$E[X] + \beta D[X].$$

As shown in Wang (1997), $E[X] + \beta D[X]$ can also be rewritten as a two-point mixture of PH-means:

$$(1 - \beta) H_1[X] + \beta H_{0.5}[X], \quad 0 < \beta < 1,$$

where $H[label]{[X]} = \int_0^x [S_{X}(t)] \, dt$ for $X \geq 0$. 
For ratemaking purposes, the right-tail deviation has some definite advantages over the standard deviation:
1. It automatically takes account of all moments without explicitly calculating them.
2. It preserves the second stochastic dominance and thus is more discriminative than the standard deviation. For layers with fixed width, the premium value decreases at higher layers, but the relative risk loading increases at higher layers; see Venter (1991).
3. It is additive for excess-of-loss layers and thus yields the same premium no matter how the risk is divided into sublayers. This is a desirable property for increased limits ratemaking.

For a layer \((a, a+h]\) of risk \(X\), the rate-on-line (ROL) is defined as the ratio of expected layer payments \(X_{(a,a+h]}\) to the limit \(h\):
\[
\text{ROL}_{(a,a+h]} = \frac{E[X_{(a,a+h]}]}{h}
\]
It can be easily verified that
\[
\lim_{h \to 0^+} \text{ROL}_{(a,a+h]} = S_X(a).
\]

In practice, the reinsurer can estimate the empirical rate-on-line using the ratio of the average observed layer payment to the limit. An empirical rate-on-line function over different layers serves as an approximation of the ddf \(S_X\).

As a pragmatic method for deciding risk loads, some reinsurers (for example Swiss Re) use a multiple of the square root of the empirical or estimated ROL. As the layers \((h\) becomes small\) are refined, this pragmatic method yields an approximation of a theoretical formula:
\[
(1 + \beta)E[X_{(a,a+h]}] + \beta D[X_{(a,a+h]}].
\]

### 10. A Bridge to Measuring Parameter Risk

Many actuaries agree on the importance of parameter uncertainty in pricing insurance risks. However, there is little agreement on how to measure the parameter risk. This situation is evident from the recent discussions on risk loads by fellows of the Casualty Actuarial Society, for example, Miccolis (1977), Feldblum (1990), and Meyers (1991). In this section we show that the right-tail deviation can be used as a bridge to modeling parameter uncertainties.

Assume that we have a finite sample of \(n\) observations from a class of identical insurance policies. An empirical estimate for the loss distribution is
\[
\hat{S}(t) = \frac{\# \text{ of observations} > t}{n}, \quad t \geq 0.
\]
Let \(S(t)\) represent the true underlying loss distribution, which is generally unknown and different from the empirical distribution \(\hat{S}(t)\).

From statistical estimation theory (for example, Lawless 1982, p. 402), for some specified value of \(t\), we can treat the quantity \([\hat{S}(t) - S(t)]/\sigma(\hat{S}(t))\) as having a standard normal distribution for large values of \(n\), where
\[
\sigma(\hat{S}(t)) = \frac{\sqrt{S(t)[1 - \hat{S}(t)]}}{\sqrt{n}}.
\]

The 100\(\eta\)% upper confidence limit for the true underlying distribution \(S(t)\) can be approximated by
\[
S^u(t) = \hat{S}(t) + \beta \sqrt{\hat{S}(t)} \left[ \frac{\hat{S}(t)}{S(t)} \right]^2, \quad \beta = \frac{q_{\eta}}{\sqrt{n}},
\]
where \(q_{\eta}\) is a quantile of the standard normal distribution: \(\Pr\{N(0,1) \leq q_{\eta}\} = \eta\).

As a means of dealing with parameter risk regarding the best-estimate \(\hat{S}_X(t)\), we use
\[
\hat{S}_X(t) = \hat{S}_X(t) + \beta \left( \sqrt{\hat{S}_X(t)} - \hat{S}_X(t) \right), \quad \beta = \frac{q_{\eta}}{\sqrt{n}},
\]
as an approximation of the 100\(\eta\)% upper confidence limit \(S^u(t)\).

Note that \(\hat{S}_X(t)\) always lies within the interval \([0,1]\), while \(S^u(t)\) can be greater than one. The ddf \(\hat{S}_X(t)\) implies a risk-load formula:
\[
E[X] + \beta D[X].
\]

### 11. Conclusion

We have discussed the new concept of right-tail deviation for a random variable \(X \geq 0\):
\[
D[X] = \int_0^\infty \sqrt{\Pr\{X > t\}} \, dt - E[X].
\]
We have shown that it is a simple yet powerful measure of the right-tail risk. It outperforms the standard deviation and Gini mean and has a great potential for actuarial applications.
ACKNOWLEDGMENT

I wish to thank John A. Beekman and Virginia R. Young for numerous helpful comments and suggestions. I also thank Paul Embrechts for pointing out a couple of useful references.

REFERENCES


APPENDIX

EXTENSION TO REAL-VALUED RANDOM VARIABLES

For a real-valued random variable $X$, we have

$$E[X] = \int_{-\infty}^{0} [S_x(t) - 1] \, dt + \int_{0}^{\infty} S_x(t) \, dt.$$ 

In terms of Mallows metric, the right-tail deviation can be extended to real-valued random variables as:

$$D[X] = M(S_x, \sqrt{S_x}) = \int_{-\infty}^{\infty} [\sqrt{S_x(t)} - S_x(t)] \, dt.$$ 

In the context of insurance assets, the main concern is the downside risk, that is, poorer than expected performance in the investment portfolios or failure of promised payments by the borrowers. If one looks at the distribution for investment prospects, the left-tail constitutes the major investment risk. As a measure of the left-tail risk, we define the following quantity:

$$D^*[X] = D[-X] = \int_{-\infty}^{\infty} [\sqrt{F_x(t)} - F_x(t)] \, dt.$$ 

Now we also define the two-sided deviation as

$$\Delta[X] = \frac{D[X] + D^*[X]}{2} = \frac{1}{2} \left\{ \int_{-\infty}^{\infty} [\sqrt{S_x(t)} + \sqrt{1 - S_x(t)} - 1] \, dt \right\},$$

and the two-sided tail index as

$$\delta(X) = \frac{\Delta[X]}{E[X]}.$$ 

Note that the two-sided deviation can be readily expressed in terms of Mallows metric

$$\Delta[X] = \frac{1}{2} M(\sqrt{S_x}, 1 - \sqrt{1 - S_x}).$$

Proposition 12.1

- $\Delta[X + b] = \Delta[X]$, for $-\infty < b < \infty$.
- $\Delta[aX] = a\Delta[X]$, for $-\infty < a < \infty$.
- $\Delta[X + Y] \leq \Delta[X] + \Delta[Y]$.
- If $X$ and $Y$ are comonotonic, then $\Delta[X + Y] = \Delta[X] + \Delta[Y]$.

Remarks

The term $\Delta[X]$ may be difficult to evaluate numerically for many parametric distributions, but there is no such problem for empirical sample distributions.

Discussions on this paper can be submitted until October 1, 1998. The author reserves the right to reply to any discussion. See the Submission Guidelines for Authors for detailed instructions on the submission of discussions.