Estimation of Distortion Risk Measures

Hideatsu Tsukahara

Department of Economics, Seijo University, Tokyo

e-mail address: `tsukahar@seijo.ac.jp`

The concept of coherent risk measure was introduced in Artzner et al. (1999). They listed some properties, called axioms of 'coherence', that any good risk measure should possess, and studied the (non-)coherence of widely-used risk measure such as Value-at-Risk (VaR) and expected shortfall (also known as tail conditional expectation or tail VaR). Kusuoka (2001) introduced two additional axioms called law invariance and comonotonic additivity, and proved that the class of coherent risk measures satisfying these two axioms coincides with the class of *distortion risk measures* with convex distortions.

To be more specific, let X be a random variable representing a loss of some financial position, and let $F(x) := \mathrm{P}(X \le x)$ be the distribution function (df) of X. We denote its quantile function by $F^{-1}(u) := \inf\{x \in \mathbb{R} : F_X(x) \ge u\}$, $0 < u < 1$. A distortion risk measure is then of the following form

$$\rho_D(X) := \int_{[0,1]} F^{-1}(u)\, \mathrm{d}D(u) = \int_{\mathbb{R}} x\, \mathrm{d}D \circ F(x), \tag{1}$$

where D is a distortion function, which is simply a df D on $[0,1]$; i.e., a right-continuous, increasing function on $[0,1]$ satisfying $D(0) = 0$ and $D(1) = 1$. For $\rho_D(X)$ to be coherent, D must be convex, which we assume throughout this paper. The celebrated VaR can be written of the form (1), but with non-convex D; this implies that the VaR is not coherent. Also note that different authors use different names *spectral risk measure* or *weighted V@R* for a distortion risk measure.

The most well-known example of coherent risk measure is the above-mentioned expected shortfall. Taking distortion of the form $D_\alpha^{\mathrm{ES}}(u) = \alpha^{-1}\big[u - (1-\alpha)\big]_+$, $0 < \alpha < 1$ yields the expected shortfall as a distortion risk measure:

$$\mathrm{ES}_\alpha(X) := \frac{1}{\alpha} \int_{1-\alpha}^{1} F^{-1}(u)\, \mathrm{d}u.$$

The following one-parameter families of distortion yields several classes of coherent risk measures:

- *Proportional hazards (PH) distortion*: $D_\theta^{\mathrm{PH}}(u) = 1 - (1-u)^\theta$,
- *Proportional odds (PO) distortion*: $D_\theta^{\mathrm{PO}}(u) = \theta u / [1 - (1-\theta)u]$
- *Gaussian distortion*: $D_\theta^{\mathrm{GA}}(u) = \Phi(\Phi^{-1}(u) + \log\theta)$

To implement the risk management/regulatory procedure using risk measures, it is necessary to statistically estimate their values based on data. For a distortion risk measure, its form (1) suggests a natural estimator which is a simple form of an L-statistic. The main theme of this paper is to derive the asymptotic statistical properties of simple estimators of those risk measures based on strictly stationary sequences, and to compare some distortion risk measures and VaR.

Let $(X_n)_{n\in\mathbb{N}}$ be a strictly stationary process with a stationary distribution F, and denote by $\mathbb{F}_n$ the empirical df based on the sample $X_1,\ldots,X_n$. A natural estimator of $\rho(X)$ is given by

$$\widehat{\rho}_n = \int_0^1 \mathbb{F}_n^{-1}(u)\,\mathrm{d}D(u) = \sum_{i=1}^n c_{ni}X_{n:i},\tag{2}$$

where $c_{ni} := D(i/n) - D((i-1)/n)$ and $X_{n:1} \le X_{n:2} \le \cdots \le X_{n:n}$ are the order statistics based on the sample $X_1,\ldots,X_n$.

In what follows, instead of restricting ourselves to the particular form (2) of L-statistic, we consider a general L-statistic of the following form:

$$T_n = \frac{1}{n}\sum_{i=1}^n c_{ni}h(X_{n:i}),\tag{3}$$

where c_{ni}'s are constants. Define for $0 \le u \le 1$,

$$J_n(u) := \sum_{i=1}^n c_{ni}\mathbf{1}_{((i-1)/n,\,i/n]}(u) + c_{n1}\mathbf{1}_{\{0\}}(u), \quad \Psi_n(u) := \int_{1/2}^u J_n(v)\,\mathrm{d}v$$

Then we have

$$T_n = \int_0^1 h(\mathbb{F}_n^{-1}(u))J_n(u)\,\mathrm{d}u = \int_{[0,1]} h(\mathbb{F}_n^{-1}(u))\,\mathrm{d}\Psi_n(u).$$

Let $g := h \circ F^{-1}$, and define the centering constants

$$\mu_n := \int_0^1 g(u)J_n(u)\,\mathrm{d}u = \int_{[0,1]} g(u)\,\mathrm{d}\Psi_n(u).$$

Consistency is a basic desirable property of statistical estimators. The following result was proved in van Zwet (1980) for the i.i.d. case, but his proof remains to be valid for the ergodic case.

Proposition 1 *Suppose that $X_1, X_2,\ldots$ forms an ergodic stationary sequence. Let $1 \le p \le \infty$, $1/p + 1/q = 1$, and assume that $J_n \in L^p(0,1)$ for $n = 1, 2,\ldots$, and $g \in L^q(0,1)$. If either*

(i) $1 < p \le \infty$ and $\sup_n \mathrm{E}(|J_n|^p) < \infty$, or

(ii) $p = 1$ and $\{J_n,\ n = 1, 2,\ldots\}$ is uniformly integrable,

then we have $T_n - \mu_n \to 0$, a.s..

Further, if there exists a function $J \in L_p$ such that $\lim_{n\to\infty}\int_0^t J_n(s)\,\mathrm{d}s = \int_0^t J(s)\,\mathrm{d}s$ for every $t \in (0,1)$, then $T_n \to \int_{[0,1]} J(s)\,\mathrm{d}g(s)$, a.s. By this result, in particular, our estimator $\widehat{\rho}_n$ in (2) of distortion risk measure proves to possess strong consistency under the very general conditions stated above.

For the asymptotic normality, we basically draw upon Shorack and Wellner (1986), Chapter 19, for the form of assumptions and the line of argument. First we set out the following assumption on (X_n).

2

(A.1) $(X_n)_{n\in\mathbb{N}}$ is strongly mixing: Setting $\mathscr{F}_i^j := \sigma(X_i,\ldots,X_j)$, the strong mixing coefficient

$$\alpha(n) := \sup\left\{|P(A\cap B)-P(A)P(B)|\colon A\in\mathscr{F}_1^k, B\in\mathscr{F}_{k+n}^\infty, k\geq 1\right\}$$

converges to 0 in such a way that

$$\alpha(n) = O(n^{-\theta-\eta}) \quad \text{for some } \theta\geq 1+\sqrt{2},\ \eta>0$$

Note that strong mixing assumption is the weakest requirement among various mixing concepts. Next we assume the bounded growth of g and J_n, and smoothness of J_n.

(A.2) h is a function of bounded variation: $h=h_1-h_2$, where h_1 and h_2 are increasing, left-continuous, and satisfy

$$|h_i(F^{-1}(u))| \leq H(u), \quad \text{for all } 0<u<1,$$

where $H(u) := Mu^{-d_1}(1-u)^{-d_2}$.

For $g = h\circ F^{-1}$, let $\int \mathrm{d}g$ be the integral with respect to the Lebesgue-Stieltjes signed measure associated with g, and $\int \mathrm{d}|g|$ be the integral with respect to the total variation measure associated with g.

(A.3) There exists a function J which is $|g|$-a.e. continuous such that J_n converges to J locally uniformly $|g|$-a.e.

(A.4) For $B(u) := Mu^{-b_1}(1-u)^{-b_2}$, $|J_n(u)|\leq B(u)$, $|J(u)|\leq B(u)$ for all $0<u<1$ with $b_1\vee b_2 < 1$.

We note that under (A.2) and (A.4),

$$\int_0^1 [u(1-u)]^r B(u)\,\mathrm{d}|g|(u) < \infty \tag{4}$$

when $r > (b_1+d_1)\vee(b_2+d_2)$ (see Shorack and Wellner (1986), Lemma 19.1.1).

Before we state and prove the asymptotic normality of the estimator (2), let us note that it is possible to reduce the argument to the uniform case, as in the i.i.d. case. Namely, there exists a strictly stationary sequence $(\xi_n)_{n\in\mathbb{N}}$ with the same mixing rate as (X_n) such that $X_n = F^{-1}(\xi_n)$ and $\xi_n\sim U(0,1)$ (on a possibly extended probability space; see Lemma 4.2 in Dehling and Philipp (2002)). Let $\mathbb{G}_n$ be the empirical df based on $\xi_1,\ldots,\xi_n$. Then

$$T_n-\mu_n \overset{\mathcal{L}}{=} = \int_{[0,1]} g(u)\,\mathrm{d}\big[\Psi_n(\mathbb{G}_n(u))-\Psi_n(u)\big]. \tag{5}$$

Here $X\overset{\mathcal{L}}{=}Y$ means that the random variables X and Y have the same distribution.

Let $C_k(u,v):=\mathrm{P}(\xi_1\leq u,\ \xi_k\leq v)$ and put

$$\sigma(u,v):= u\wedge v-uv+\sum_{k=2}^\infty[C_k(u,v)-uv]+\sum_{k=2}^\infty[C_k(v,u)-uv]. \tag{6}$$

When (ξ_n) satisfies the same mixing rate as in (A.1), it follows from the covariance inequality (see Dehling and Philipp (2002), Lemma 3.9) that the two series on the right-hand side of (6) are absolutely convergent. We define the empirical process $\mathbb{U}_n(u):=\sqrt{n}(\mathbb{G}_n(u)-u)$ as usual.

Theorem 2 *Let $(X_n)_{n \in \mathbb{N}}$ be a strictly stationary sequence satifying (A.1)–(A.4) with*

$$b_i + d_i + \frac{2b_i + 1}{2\theta} < \frac{1}{2}, \quad i = 1, 2 \tag{7}$$

Then we have

$$\sqrt{n}(T_n - \mu_n) \xrightarrow{\mathcal{L}} N(0, \sigma^2),$$

where

$$\sigma^2 := \int_0^1 \int_0^1 \sigma(u, v) J(u) J(v) \, \mathrm{d}g(u) \mathrm{d}g(v) < \infty \tag{8}$$

Returning to the problem of estimating distortion risk measures, we should set $c_{ni} = n[D(i/n) - D((i-1)/n)]$ and $h(x) = x$. Then in most cases, the limit of J_n will be d, so applying Theorem 2 we have the following corollary.

Corollary 3 *Assume (A.1), (A.2) with $h(x) = x$, (A.3) with $J = d$, and (A.4). Then, for the estimator $\widehat{\rho}_n$ of (2), we have*

$$\sqrt{n}(\widehat{\rho}_n - \rho(X)) \xrightarrow{\mathcal{L}} N(0, \sigma^2),$$

where

$$\sigma^2 = \int_0^1 \int_0^1 \sigma(u, v) d(u) d(v) \, \mathrm{d}F^{-1}(u) \mathrm{d}F^{-1}(v).$$

When we try to construct approximate confidence intervals for risk measures, we need to estimate the asymptotic variance (8). Let

$$Y_n := \int_{[X_n, \infty)} J(F(x)) \, \mathrm{d}h(x), \quad n \in \mathbb{Z}.$$

It is then easy to see that σ^2 is written as the double-sided infinite sum of autocovariance $\gamma(n)$ of the stationary sequence (Y_n). It is well known that

$$\sum_{n=-\infty}^{\infty} \gamma(n) = 2\pi f(0),$$

where f is the spectral density of γ. Thus our problem is to estimate $f(0)$, and we would use a consistent estimator of $f(0)$ as given in Brockwell and Davis (1991). But F in the expression of Y_n is unknown, so we must replace it with the empirical distribution function. That is, we use

$$Y_{i,n} := \int_{[X_i, \infty)} J(\mathbb{F}_n(x)) \, \mathrm{d}h(x), \quad i = 1, \ldots, n$$

in estimating $f(0)$. This should give a consistent estimator of the asymptotic variance (8).

Example 4 (Inverse-gamma autoregressive stochastic volatility) In order for us to be able to compute the true values of various risk measures with adequate accuracy so that we can evaluate the estimation bias and root mean squared error (RMSE), we introduce the following simple stochastic volatility model. Let $X_t = \sigma_t Z_t$ and suppose

4

that $V_t := 1/\sigma_t^2$ follows the first-order autoregressive gamma process introduced in Gaver and Lewis (1980):

$$V_t = \rho V_{t-1} + \varepsilon_t,$$

where V_t has a gamma distribution with shape parameter α and inverse-scale parameter β for each t, (ε_t) is a sequence of i.i.d. random variables, and $0 \le \rho < 1$. It is known that the distribution of ε_t is compound Poisson.

Let (Z_t) be a sequence of independent random variables with standard normal distribution, which are also independent of (ε_t). Then it is well known, especially in Bayesian analysis, that X_t has a scaled t-distribution with 2α degrees of freedom and scale papram-eter $\sigma^2 = \beta/\alpha$; this allows us to calculate the true values of VaR, expected shortfall, and proportional odds risk measure. Also (V_t) can be shown to be geometrically ergodic, so the resulting (X_t) is also geometrically ergodic, and hence exponentially strong mixing. Thus our assumption (A.1) is satisfied in this model.

To make use of the setting in Example 2.21 of McNeil et al. (2005), we chose $\alpha = 2$, $\beta = 16000$, $\rho = 0.5$, so that X_t has a scaled t-distribution with four degrees of freedom, and its standard deviation is equal to $2000/\sqrt{250} \approx 126.5$. For this case, the true values of VaR, expected shortfall, and proportional odds risk measures are given in Table 4.1 of Tsukahara (2009). For $\theta = 0.1$, 0.05, 0.01, we generated 1000 samples of size 500 and computed the estimates, the estimated biases, and the RMSEs for our estimator. For the purpose of comparison, we also perform the same procedure with i.i.d. observations from a scaled t-distribution with four degrees of freedom. The results are summarized in Table 1.

Table 1: Simulation results for estimating VaR, ES and PO risk measures with inverse-gamma autoregressive SV observations with $t(4)$ marginal and with i.i.d. $t(4)$ observations ($n = 500$, # of replication = 1000)

	$\theta = \alpha$	VaR bias	VaR RMSE	ES bias	ES RMSE	PO bias	PO RMSE
	0.1	0.0692	10.9303	−2.2629	22.1361	−1.7739	17.5522
SV	0.05	2.5666	17.6755	−1.2168	37.2719	−2.0200	28.5053
	0.01	14.9577	61.2290	−11.9600	103.9269	−15.7888	73.7147
	0.1	0.7976	10.5893	−1.2914	19.5756	−1.3574	15.3271
i.i.d.	0.05	0.7974	16.1815	−2.6346	31.3166	−2.8342	23.9933
	0.01	10.6838	53.2567	−12.9355	95.9070	−15.8086	69.5425

They show clearly that both biases and RMSEs increase for all three risk measures as θ gets smaller; this is expected from the asymptotic results. Hence estimation with small θ is a difficult task even with moderate sample size of $n = 500$. Maybe this shows the limitation of purely statistical methods for estimating the values of risk measures.

The estimated RMSEs are large probably reflecting the heavy tail of the t-distribution with four degrees of freedom. Although RMSE is slightly smaller for every risk measure in the i.i.d. case, there does not seem to be a big difference in the behavior of the estimates between in the stochastic volatility case and i.i.d. case, reflecting perhaps the quite weak

5

dependence in this stochastic volatility model.

Note that systematic negative biases are observed for our L-statistics type estimators in cases of expected shortfall and proportional odds risk measures. Examining the histograms (not shown here) shows that the distribution of the estimator is right-skewed with this sample size. We suggest that some kind of bias reduction method be applied in practice.

REFERENCES

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance*, **9**, 203–228.

Brockwell, P. J. and R. A. Davis (1991). *Time Series: Theory and Methods*, New York: Springer-Verlag.

Dehling, H. and W. Philipp (2002). Empirical process techniques for dependent data. In H. Dehling, T. Mikosch and M. Sørensen (eds.), *Empirical process techniques for dependent data*, Boston: Birkhäuser, p.p. 3–113.

Gaver, D. P. and P. A. W. Lewis (1980). First-order autoregressive gamma sequences and point processes, *Advances in Applied Probability*, **12**, 727–745.

Kusuoka, S. (2001). On law invariant coherent risk measures, *Advances in Mathematical Economics*, **3**, 83–95.

McNeil, A. J., R. Frey, and P. Embrechts (2005). *Quantitative Risk Management: Concepts, Techniques, and Tools*, Princeton, New Jersey: Princeton University Press.

Shorack, G. R. and J. A. Wellner (1986). *Empirical Processes with Applications to Statistics*, New York: John Wiley & Sons.

Tsukahara, H. (2009). One-parameter families of distortion risk measures, *Mathematical Finance*, **19**, 691–705.

van Zwet, W. R. (1980). A strong law for linear functions of order statistics, *Annals of Probability*, **8**, 986–990.