

Risk tendency measures for ordinal variables

Paolo Giudici and Silvia Figini

Abstract In this paper we present a novel index useful to derive a measure of risk on the basis of ordinal variables. The content of this work is related to coherent measures of risk proposed in the financial literature. We point out that our proposal is interesting also for specific applications in non financial sectors. Empirical evidences are given on a real case study.

Key words: Risk integration, Ordinal variables, Coherent measures of risk

1 Introduction

The objectives of this paper are: to review the definition of coherent measure of risk expressed for continuous risk variables and to propose a novel index of risk for ordinal variables.

The paper is structured as follows: Section 2 reviews the axioms of coherent measures of risk with specific attention on sub-additivity; Section 3 shows our proposal and the related properties; Section 4 presents an application and, finally, Section 5 reports the conclusions and further ideas of research.

2 Coherent risk measures for continuous variables

In the field of financial economics there are a number of ways that risk can be defined; to clarify the concept theoreticians have described a number of properties

Paolo Giudici
University of Pavia, e-mail: giudici@unipv.it

Silvia Figini
University of Pavia e-mail: silvia.figini@unipv.it

that a risk measure might have.

Following Artzner et al. 1999, a risk measure ρ is coherent if the following four axioms hold:

- **Sub-additivity:** $\rho(X + Y) \leq \rho(X) + \rho(Y)$;
- **Homogeneity:** for any number $\alpha \geq 0$, $\rho(\alpha X) = \alpha \rho(X)$;
- **Monotonicity:** $\rho(Y) \geq \rho(X)$ if $X \leq Y$;
- **Risk free condition:** $\rho(X + k) = \rho(X) - k$, for any constant k .

Coherent measures of risk are relevant, for example, in the financial area, where Value-at-Risk (VaR) has been widely accepted as a tool for risk management.

Following the guidelines of the Basel Committee, financial regulators all over the world have adopted VaR for designing capital adequacy standard for banks and financial institutions. Apart from financial regulators like central banks and Securities Exchange regulators, financial firms have adopted VaR for internal risk management and allocation of resources.

Despite its global recognition as a measure of financial risk, VaR has been criticised on certain theoretical grounds for not satisfying the conditions of a coherent risk measure Wilmott (2006) and Artzner et al. (1999).

More precisely, VaR being a tail-quantile, is not sub-additive in general, and therefore, is not coherent. Furthermore, VaR is sub-additive under the restrictive assumption of normally distributed asset returns, which is a rare situation in practical financial applications.

Let X and Y be two financial assets. As pointed out before, a risk measure ρ is sub-additive if the following condition is true:

$$\rho(X + Y) \leq \rho(X) + \rho(Y) \quad (1)$$

Thus, the risk measure of the sum of two assets is bounded above by the sum of their individual risks.

In our opinion, the property of sub-additivity can be motivated by many practical considerations: it ensures that the diversification principle of modern portfolio theory holds and we remark that a sub-additive measure would always generate a lower risk measure for a diversified portfolio than a non-diversified portfolio.

In terms of internal risk management, sub-additivity also implies that the overall risk of a financial firm can be added up to be equal to or less than the sum of the risks of individual departments of the firm. This appears to give an appealing dimension into the idea of integrated risk management.

Our final consideration is that in the absence of subadditivity, a financial firm with risk $X + Y$ would calculate a smaller regulatory capital of $\rho(X) + \rho(Y)$, leading to underestimation of risk reserve.

It can be proved easily by constructing suitable examples that VaR violates sub-additivity property (Artzner et al., 1999; Acerbi et al., 2001).

In order to demonstrate that VaR violates sub-additivity, we propose a simple example.

Consider two assets X and Y that are usually normally distributed, but subject to the

occasional independent shocks: $X = \varepsilon + \eta$, $\varepsilon \sim N(0, 1)$, where $\eta = 0$ with probability 0.991 or $\eta = -10$ with probability 0.009.

In this case the 1% VaR is 3.1, since the probability that η is -10 is less than 1%. Suppose that Y has the same distribution independently from X , and we formulate an equally weighted portfolio of X and Y . In that case, the 1% portfolio VaR is 9.8, because for $(X + Y)$ the probability of getting the -10 draw for either X or Y is much higher.

$$VaR(X + Y) = 9.8 > VaR(X) + VaR(Y) = 3.1 + 3.1 = 6.2$$

The lack of sub-additivity of VaR has been severely criticised, and several alternative measures have been proposed to replace VaR such as tail conditional expectation (TCE) and worst conditional expectation (WCE) by Artzner et al. (1999) and expected shortfall (ES) by Acerbi et al. (2001). It is however, argued that imposing sub-additivity for all risks (including dependent risks) is not in line with what could be called best practice (Dhaene et al., 2003).

In order to relax the assumptions for the VaR computation, in this paper we propose a novel measure of risk. Our index could be applied in a wide range of statistical measures of losses without require a specific parametric assumption for the loss distribution.

For example, focusing on operational risk management the distribution that we have available at the ordinal scale is the convolution between the frequencies of the events (frequency distribution) and the relative severity (severity distribution), as described in more details in Section 3.

3 Coherent risk measures for ordinal variables

As pointed out in Section 2, in this section we explain how to derive a measure of risk starting from one or more that one ordinal variables.

For sake of simplicity we refer to operational risk.

Operational data for risk measurement are typically summarised in a matrix composed of J business lines (rows) and I event types (columns).

Let E_{ij} the i -th event type ($i = 1, \dots, I$) in the j -th business line ($j = 1, \dots, J$).

For each event type in a specific business line, we have two different measures: the frequency (how many times the specific event type appears in the business line) and the severity, expressed in an ordinal scale (i.e. H=high, L=medium and L=low).

In order to summarise the frequency and the related severity, we may structure a data set, which counts for each event type business line and for a given severity, the absolute frequency. This procedure is repeated across the J business lines considered in the analysis.

Considering the first event type, n_{H11} is the number of times for which the first event appears with high severity in the first business line; n_{M11} is the number of times for

which the first event appears with medium severity in the first business line and n_{L11} is the number of times for which the first event appears with low severity in the first business line. In general for the i -th event in the j -th business line, let n_{Hij} , n_{Mij} and n_{Lij} the the number of times for which the event E_i , $i = 1, \dots, I$ appears in the j , $j = 1, \dots, J$ business line with high, medium or low severity respectively.

The theoretical matrix is composed of K rows, representing the business lines and $I \times 3$ columns equivalently to the number of events multiplied by the levels of severity under analysis, (high=H, medium=M and low=L), where each cell represents a mutually exclusive combination of business line (row) and event type (column) values. N represents the total number of events registered in the K business lines for the specified level of severity.

Our proposal improves the state of the art in operational risk management because it derives on the basis of ordinal variables related to I event types and K business lines a measure of risk.

The measure proposed is based on the cumulative distribution function. It is well known that the cumulative distribution function is a monotonously increasing (or more precisely, non decreasing) function bounded between 0 and 1.

We derive for each event and for a given business line, the cumulative relative frequency distribution, starting from a generic element p_{SKI} , where S is the severity level (H=high, M=medium, L=low), K is the k -th business line and I is the i -th event.

Let F_{ski} be the cumulative distribution function calculated in such point; we propose to employ the following index as a measure of risk:

$$I_{ik} = \sum_{s=1}^S F_{sij}, \quad i = 1, \dots, I, \quad k = 1, \dots, K, \quad s = 1, \dots, S, \quad (2)$$

where $F_{ij} = \sum_{s=1}^S p_{ski}$, $s = 1, 2, 3$, where $s = 1$, $s = 2$ and $s = 3$ is the lower, medium and high level of severity respectively.

We remark that I takes values in $[0, \infty)$ and more precisely $I = 0$ when a specific event in a specific business lines never appear with severity equal to 1, 2 or 3.

The normalised index, I^* is bounded between 0 and 1. More precisely $I^* = 0$ when for the specific business lines the event type under consideration never appears, while $I^* = 1/S$ when the frequencies of the event type in the specific business line are equal for the levels low, medium and high of severity.

From a methodological viewpoint, I^* as pointed out before, is based on the cumulative distribution function and more precisely two measures of risks related to two different event types differ only for their relative frequencies collected in correspondence to the severity levels considered.

We remark that the final risk estimation does not depend on the normalisation, but is directly linked to the cumulative distribution function. This leads us to compare the relative measures of risk (I^*) through stochastic dominance, as shown in Figini et al. 2010.

4 Application

In order to show how our proposal works, we employ a real data set provided by Tadiran, a partner in the european project Musing.

The data set we use is composed of 1126 statistical units (PBX). The data is collected by the Tadiran call center. A customer of a specific business line (i.e. Banking, Health, Defence, Cooperative etc) calls Tadiran to signal specific problems about hardware (H), software (ST), interface (I), network communications (NC) or security (S). The call center operator inputs to the system a the reference to the call listing the PBX number and the level of severity of the problem as reported by the customer (high, medium, low).

The data structure of the data collected is a data base with problem reports listed by PBX. For each problem, the operator enters a level of severity (low=1, medium=2, high=3). A single customer could reports more then one problem in a time window. For example a customer in the banking business line called Tadiran three times to report problems of hardware (severity=1), software (severity=2), interface (severity=1). It is possible also that a customer in the same call reports to more that one problem.

Table 1 reports for each business line the number of events registered and the number of statistical units. We have used the results summarised for specific problems

Business line	PBX	Total Events
Banking	86	222
Finance	28	200
Health	33	171
Industry	20	132
General	28	119
Cooperatives	31	107
Municipalities	21	95
Hotels	22	80
Leisure Consumerism	8	70
Transportation	9	67
Computers	4	43
Government owned company	9	39
Electronics	6	35
Defense	11	29
Education	5	11
Elderly Citizens home	4	6
Construction	1	4
Government	2	4

Table 1 Business lines and related events

(hardware, software and security), grouped in specific business lines to build our index. The results are reported in Table 2. As we can observe in Table 2 for the hardware problem, the highest risk is for Health and Banking, while the lowest risk

BL	Total events	Hardware	Software	Security
Banking	222	0.12	0.42	0.51
Finance	200	0.08	0.38	0.37
Health	171	0.16	0.36	0.24
Industry	132	0.01	0.22	0.19
General	119	0.08	0.17	0.25
Cooperatives	107	0.03	0.23	0.14
Municipalities	95	0.06	0.20	0.12
Hotels	80	0.02	0.11	0.07
Leisure Consumerism	70	0.04	0.19	0.06

Table 2 Index for specific problems

is for Industry. Considering software problem, the highest risk is for Banking and finance, while the lowest risk is for Hotels. Finally, looking at the security problem, the highest risk is for Banking and finance, while the lowest risk is for Leisure Consumerism. As we observe from Table 2, our index is able to give a simple measure bounded between 0 and 1 which highlights specific alarms in the business lines considered in the analysis. This information is useful in order to prioritise particular interventions and to plan maintenance.

5 Conclusions

In this paper we have reviewed coherent measures for risk for continuous variables. As novelty, we have proposed an index useful to measure risk on the basis of ordinal variables. Empirical evidences are achieved on a real data set. We think that our methodological approach could be interesting also for different applications, such as credit risk. Further ideas of research will consider integration and stochastic dominance ordering among the risk measures derived.

Acknowledgements The authors thank the European Project Musing for the financial support.

References

1. Acerbi, C. and Tasche, D.: On the Coherence of Expected Shortfall, working paper, available at <http://www.gloriamundi.org/var/wps.html> (2001).
2. Artzner, P., Delbaen, F., Eber, J.M. and Heath, D.: Coherent Measures of Risk, *Mathematical Finance* 9 vol. 3, 203-228 (1999).
3. Figini, S. and Giudici, P.: Risk measures for ordinal variables, submitted (2010).
4. Figini, S., Salini, S. and Kenett, R.: Integrating Operational and Financial Risk Assessments, submitted (2010).
5. Wilmott P.: *Quantitative Finance*, Wiley (2006).