PROJECT ALETHEA — RESEARCH MONOGRAPH SERIES (PAPER 1.1 EXTENDED)

Axiomatic Superiority of Expected Shortfall over Value at Risk in Autonomous Risk Verification

Author: Ying Jack Chin (Chief Risk Officer) **Published:** November 2026 **Hub:** aegislab.io

1. Abstract & Executive Summary

In high-dimensional quantitative risk deployment and autonomous agent execution, failure to bound tail risk introduces catastrophic systemic vulnerability. While **Value at Risk (VaR_α)** remains embedded in legacy regulatory architecture, its failure to satisfy the subadditivity axiom under non-Gaussian loss distributions creates hazardous incentive structures for risk managers and continuous neural architectures. This monograph expands upon Section 3 of *Project Alethea Paper 1*, evaluating the measure-theoretic foundations of risk measures across non-convex probability topologies.

We prove why **Expected Shortfall (ES_α)**—and its generalization into Spectral Risk Measures—serves as a strictly superior governance invariant for institutional capital allocation and autonomous AI verification pipelines. Furthermore, we address the elicitability paradox (Fissler & Ziegel, 2016), illustrating how joint scoring functions permit rigorous backtesting of ES_α while retaining strict subadditivity within Satisfiability Modulo Theories (SMT) verification loops.

2. Measure-Theoretic Formalization: VaR_α vs. ES_α

Let $(\Omega, F, \mathbb{P})$ be an non-atomless probability space representing financial market states or state-action transition spaces in autonomous execution engines. Let $X \in L^1(\Omega, F, \mathbb{P})$ be a real-valued random variable representing net loss over horizon T, where positive values denote financial loss or safety invariant violations.

2.1 Value at Risk (VaR_α)

For a confidence level $\alpha \in (0,1)$ (typically $\alpha = 0.99$ or 0.995), $VaR_\alpha(X)$ measures the lower α-quantile of the loss distribution:

> $$\mathbf{VaR_\alpha(X)} = \inf\{ l \in \mathbb{R} : \mathbb{P}(X > l) \leq 1 - \alpha \} = \inf\{ l \in \mathbb{R} : F_X(l) \geq \alpha \}$$

where $F_X(l) = \mathbb{P}(X \leq l)$ is the cumulative distribution function (CDF) of loss X.

Structural Limitation: VaR_α reports strictly the point threshold loss at quantile α. It is fundamentally blind to the magnitude, skewness, and heavy-tailed properties of the loss distribution beyond α (i.e., $\mathbb{E}[X \mid X > VaR_\alpha]$). Two asset portfolios with identical 99% VaR can exhibit orders of magnitude difference in expected crash severity.

2.2 Expected Shortfall (ES_α)

To evaluate tail severity, Expected Shortfall (Conditional Value at Risk, CVaR, or Tail VaR) evaluates the conditional expectation of loss exceeding VaR_α. For continuous loss distributions, $ES_\alpha(X)$ is defined as:

> $$\mathbf{ES_\alpha(X)} = (1 / (1 - \alpha)) \int_\alpha^1 \mathrm{VaR}_u(X)\, du = \mathbb{E}\,[\, X \mid X \geq \mathrm{VaR}_\alpha(X)\,]$$

When F_X contains discontinuities or point masses (common in credit default jumps, illiquid order books, and algorithmic liquidation cascades), generalized quantile definitions are required to preserve continuity and coherence:

> $$\mathbf{ES_\alpha(X)} = (1 / (1 - \alpha))\, (\, \mathbb{E}[\, X \cdot \mathbb{1}_{\{X \geq \mathrm{VaR}_\alpha(X)\}}\,] + \mathrm{VaR}_\alpha(X)\, (\, \mathbb{P}(X < \mathrm{VaR}_\alpha(X)) - \alpha\,)\,)$$

3. The Artzner Coherence Axioms & Mathematical Proofs

Artzner, Delbaen, Eber, and Heath (1999) established that a mapping $\rho: L^\infty(\Omega) \to \mathbb{R}$ is a **coherent risk measure** if and only if it satisfies four fundamental postulates:

- **1. Subadditivity:** $\rho(X + Y) \leq \rho(X) + \rho(Y)$ for all $X, Y \in L^\infty$
- **2. Monotonicity:** If $X \leq Y$ almost surely, then $\rho(X) \leq \rho(Y)$
- **3. Positive Homogeneity:** $\rho(\lambda X) = \lambda \rho(X)$ for all $\lambda \geq 0$
- **4. Translation Invariance:** $\rho(X + c) = \rho(X) + c$ for all $c \in \mathbb{R}$

3.1 Failure of Subadditivity in VaR_α

VaR_α **fails the subadditivity axiom** under fat-tailed, skewed, or discrete loss distributions. That is, there exist assets X and Y such that $VaR_\alpha(X + Y) > VaR_\alpha(X) + VaR_\alpha(Y)$.

> **Mathematical Counterexample (Tail Concentration Risk)**
> Consider two independent digital default options X and Y. Each option pays a loss of \$100,000 with probability 0.04, and \$0 with probability 0.96.
> • Evaluating $VaR_{0.95}(X)$: Since $\mathbb{P}(X > 0) = 0.04 < 0.05$, $VaR_{0.95}(X) = \$0$.
> • Similarly, $VaR_{0.95}(Y) = \$0$. Sum of individual risk measures: $VaR_{0.95}(X) + VaR_{0.95}(Y) = \$0$.
> • Combine the portfolio $Z = X + Y$. The probability of at least one default is $\mathbb{P}(Z > 0) = 1 - (0.96 \times 0.96) = 0.0784$ (7.84%).
> • Because 7.84% > 5%, the 95th percentile loss of the combined portfolio Z yields $VaR_{0.95}(X + Y) = \$100{,}000 > \$0 + \$0$.
> **Governance Implication:** Under VaR, merging independent risks inflates reported risk, while splitting high-risk assets across subsidiaries artificially suppresses reported capital requirements—creating structural systemic risk.

3.2 Complete Proof of Coherence for ES_α

Unlike VaR_α, ES_α strictly satisfies all four axioms. Subadditivity follows from Kusuoka's characterization and the dual representation theorem over scenario sets:

$$\mathbf{ES_\alpha(X)} = \sup \{ \mathbb{E}_\mathbb{Q}[X] : \mathbb{Q} \ll \mathbb{P}, d\mathbb{Q}/d\mathbb{P} \le (1 / (1 - \alpha)) \}$$

Let P_α be the set of density generators $L = d\mathbb{Q}/d\mathbb{P}$ satisfying $0 \le L \le 1/(1-\alpha)$ and $\mathbb{E}[L] = 1$. For any random variables X and Y:

$$\mathrm{ES}_\alpha(X + Y) = \sup_{L \in P_\alpha} \mathbb{E}\,[\,(X + Y)\,L\,] \le \sup_{L \in P_\alpha} \mathbb{E}[\,X\,L\,] + \sup_{L \in P_\alpha} \mathbb{E}[\,Y\,L\,] = \mathrm{ES}_\alpha(X) + \mathrm{ES}_\alpha(Y)$$

Because the supremum over linear operators is subadditive, ES_α preserves subadditivity across all probability spaces and loss distributions without exception.

4. Spectral Risk Measures & Kusuoka Representation

Expected Shortfall is a specific instance of the broader class of **Spectral Risk Measures** (Acerbi, 2002). A spectral risk measure $M_\phi(X)$ applies a subjective risk-aversion spectrum weighting function $\phi(u)$ to quantiles:

$$\mathbf{M_\phi(X)} = \int_0^1 \phi(u)\,\mathrm{VaR}_u(X)\,du$$

where $\phi(u)$ satisfies: (1) $\phi(u) \ge 0$, (2) $\int_0^1 \phi(u)\,du = 1$, and (3) $\phi(u)$ is non-decreasing ($\phi'(u) \ge 0$). The non-decreasing property ensures that higher losses receive greater weights, guaranteeing subadditivity.

For Expected Shortfall, the weighting profile takes the binary step form: $\phi(u) = (1 / (1 - \alpha))\ \mathbb{1}_{\{u \ge \alpha\}}$. Under Kusuoka's Theorem, every law-invariant coherent risk measure on L^∞ can be represented as the supremum of spectral risk measures, proving that ES_α constitutes the foundational building block for all mathematically sound risk architectures.

5. The Elicitability Paradox & Joint Backtesting

A functional statistic $T(F)$ is **elicitability-admissible** if it minimizes the expected value of a strictly consistent scoring function $S(y, x)$:

$$\mathrm{T(F)} = \mathrm{argmin}_x \int \mathrm{S(y,\,x)}\,\mathrm{dF(y)}$$

Gneiting (2011) proved that VaR_α is elicitable using the pinball loss function, whereas ES_α **is not individually elicitable**. This led legacy practitioners to argue that Expected Shortfall cannot be backtested efficiently.

5.1 The Fissler-Ziegel Resolution

Fissler and Ziegel (2016) resolved this paradox by proving that Expected Shortfall is **jointly elicitable with VaR**. The vector-valued statistic $(VaR_\alpha(X), ES_\alpha(X))$ is elicitable under the joint scoring class $S(v, e, y)$:

> $\mathbf{S(v, e, y)} = (\mathbb{1}_{\{y \le v\}} - \alpha) g_1(v) - \mathbb{1}_{\{y \le v\}} g_1(y) + g_2(e) (e - v + (1/(1-\alpha))(v - y) \mathbb{1}_{\{y \le v\}}) - G_2(e)$

where $G_2' = g_2$ and g_2 is strictly increasing. This mathematical breakthrough allows institutional quantitative desks and automated verifiers to backtest and calibrate ES_α models with full statistical precision.

6. Institutional Comparison & Convex Optimization

Risk Dimension	Value at Risk (VaR_α)	Expected Shortfall (ES_α)
Mathematical Coherence	Non-Coherent (Violates Subadditivity)	**Coherent** (Satisfies all 4 axioms)
Tail Severity Sensitivity	Completely Blind (Binary Cutoff)	**Full Expectation Capture**
Optimization Topology	Non-Convex (Multiple Local Minima)	**Convex** (Linear Programming Solvable)
Backtesting & Elicitability	Individually Elicitable (Pinball Loss)	**Jointly Elicitable with VaR** (Fissler-Ziegel)
Regulatory Standard	Legacy Basel II / Basel III	**Basel IV (FRTB Trading Book)**

6.1 Portfolio Optimization & Linear Formulations

When optimizing portfolio allocations $w \in \mathbb{R}^n$ under VaR_α, the objective function produces a non-smooth, non-convex landscape with multiple local minima, making gradient-based solvers fail. Rockafellar and Uryasev (2000) demonstrated that ES_α transforms into a global linear programming problem over historical or Monte Carlo scenarios y_k $(k = 1, ..., S)$:

> $\mathbf{\min_{(w, a, z)} F_\alpha(w, a)} = a + (1 / (S(1 - \alpha))) \sum_{k=1}^{S} z_k$
> *subject to:* $z_k \ge f(w, y_k) - a, \quad z_k \ge 0, \quad w \in W$

This formulation guarantees efficient, unique convergence even across millions of high-dimensional constraints.

7. Bounding Non-Deterministic AI in Project Alethea

In autonomous AI execution—such as LLM agentic financial routing, automated credit underwriting, and continuous deep reinforcement learning—probabilistic risk metrics fail for identical reasons. Continuous neural networks optimize expected loss over training distributions. When exposed to out-of-distribution (OOD) tail events, an unconstrained LLM agent evaluated purely on median probabilistic bounds (VaR) can exhibit catastrophic failure modes while technically satisfying threshold checks.

Within **Project Alethea**, we enforce ES_α as a symbolic verification invariant within SMT (Satisfiability Modulo Theories) solver constraints (e.g., Z3 / CVC5):

> **SMT Invariant Constraint:**
> $\forall x \in \text{StateSpace}, \ (\mathbb{E}[\text{Loss}(x) \mid \text{Loss}(x) \geq \text{Quantile}_{\alpha}] \leq K_{\text{capacity}}) \wedge (\text{SubadditivityBound}(x) \leq \text{Threshold})$

By binding the agent's action space to an Expected Shortfall invariant rather than a VaR cutoff, the symbolic verifier guarantees that the expected loss in extreme failure scenarios remains strictly bounded within institutional capital capacity limits.

8. Conclusion & Governance Roadmap

Value at Risk served as a useful historical stepping stone, but its structural blind spots—specifically non-subadditivity and tail blindness—render it dangerous for modern quantitative finance and autonomous AI safety. Expected Shortfall provides the necessary mathematical coherence, subadditivity, spectral extension, and convex optimization properties required to govern complex financial balance sheets and autonomous AI execution engines.