Article

Not peer-reviewed version

Stochastic Hardness in Model Risk: A Framework for Tangible and Auditable Assessments in Banking

Michael Rey *

Posted Date: 12 September 2025

doi: 10.20944/preprints202509.1094.v1

Keywords: model risk; stochastic complexity; computational hardness; banking regulation; quantitative finance; tail risk

Article

Stochastic Hardness in Model Risk: A Framework for Tangible and Auditable Assessments in Banking

Michael Rey

Octonion Group, Hong Kong; contact@octoniongroup.com

Abstract

Banking executives need quantitative, auditable tools to assess model reliability that satisfy regulatory requirements while reducing operational overhead. This paper introduces a statistical stress-testing framework for model reliability assessment that provides explicit pass/fail thresholds for regulatory compliance. Our approach transforms subjective model validation into objective, data-driven certification through tail-exponent analysis and summability testing. The framework aims to deliver a polynomial-tail threshold at $p = 1$ as a concrete decision criterion, automated monitoring systems that integrate with existing infrastructure, and comprehensive audit trails that strengthen regulatory confidence. The methodology is designed to align with OCC 2011-12, SR 11-7, and Basel III requirements as a design objective while potentially providing superior risk detection compared to traditional backtesting approaches. We present a theoretical framework with operationalization pathways, without claiming realized financial impacts or supervisory approvals. Our choice of tail-exponent analysis is motivated by its direct link to computational complexity: when error rates decay faster than $1/n$, reliability scales polynomially with input size; slower decay indicates potential NP-hard explosion of unreliability. This makes the tail index a natural diagnostic for distinguishing tractable from intractable model behavior. We motivate this approach via the WC–SP dimension [4]: (SP = distributional Stochastic Polynomial-time; "WC" here is a shorthand for worst-case, NP-hard–like regimes) worst-case intractability (WC-like regimes) versus distributional polynomial-time behavior (SP) on operational inputs. The critical issue is whether reliability scales tractably on typical data; tail-exponent monitoring offers an auditable proxy for staying in the SP-like regime without invoking worst-case assumptions.

Keywords: model risk; stochastic complexity; computational hardness; banking regulation; quantitative finance; tail risk

---

1. Executive Summary

Traditional validation treats errors as purely statistical, but fails to distinguish between tractable and intractable failure modes. In complexity theory, NP-hard problems are characterized by super-polynomial explosion of effort. Analogously, model error rates that decay too slowly with input size indicate a regime of uncontrolled complexity. The tail index provides a simple, auditable threshold between these two worlds: above the threshold, errors remain manageable with polynomial resources; below it, reliability deteriorates like an NP-hard problem.

1.1. Business Problem and Regulatory Context

Model risk management in banking faces three critical challenges that directly impact regulatory compliance and operational efficiency. First, traditional validation methods rely heavily on subjective expert judgment, creating inconsistency across model reviews and difficulty in demonstrating objective compliance to supervisors. Second, existing approaches often fail to detect model deterioration early enough to prevent significant losses, as evidenced by recent regulatory findings where banks faced substantial penalties for inadequate model risk management. Third, the manual nature of current validation processes creates substantial operational overhead, with large banks spending hundreds of millions of dollars annually on model validation activities.

Regulatory authorities have responded with increasingly stringent requirements. OCC 2011-12 mandates comprehensive model validation including conceptual soundness review, ongoing monitoring, and outcomes analysis. SR 11-7 requires independent validation and clear governance frameworks. Basel III emphasizes the need for robust stress testing and capital adequacy assessments that depend critically on model reliability. These requirements create both compliance obligations and business opportunities for institutions that can demonstrate superior model risk management capabilities.

1.2. Solution Overview: Statistical Stress-Testing for Model Reliability

The Stochastic Hardness Framework provides a quantitative, auditable methodology for assessing model reliability that aims to address regulatory requirements while reducing operational costs. At its core, the framework performs statistical stress-testing of model error rates across different input sizes, providing an objective measure of model reliability through tail-exponent analysis.

The key innovation lies in translating complex mathematical concepts into practical tools that model risk managers can implement immediately. Rather than relying on subjective assessments, the framework provides a concrete decision criterion: models with tail exponents above 1.0 demonstrate mathematical theoretical conditions for long-term reliability, while those below 1.0 require remediation or rejection. This creates an objective, auditable standard that satisfies regulatory requirements for independent validation.

1.3. Why the WC–SP Dimension Matters (Motivation)

Worst-case vs. typical-case. Traditional model validation evaluates statistical fit but remains agnostic about *algorithmic regimes*. Worst-case complexity (WC-like) can be intractable even when average behavior on operational inputs is tractable (SP-like). Supervisors ultimately care about reliability *on the distributions a bank actually faces*.

Operational question. Are we operating in a tractable (SP-like) regime where reliability scales with polynomial resources, or drifting toward an intractable (WC-like) regime where unreliability explodes?

Why tails. The *tail index* summarizes how fast error rates drop as input complexity grows. A sufficiently light tail indicates that cumulative error remains bounded under scaling (SP-like); heavy tails signal risk of accumulation and regime shift.

Actionable threshold. Using a simple auditable threshold for the tail exponent provides an early-warning control: above the threshold, continue with standard monitoring; near or below it, escalate and remediate before the system enters an intractable regime.

Implication. Tail-based monitoring is not a replacement for conceptual soundness or backtesting; it complements them by adding a *complexity-awareness layer* that is distributionally grounded.

Terminology (Quick Labels)

SP denotes the distributional *Stochastic Polynomial-time* notion used in this paper: on the bank's operational input distributions, there exists a polynomial-time procedure whose per-length error rates are summable [5] (cf. Section 4). The label "*WC*" is used *informally* here as a compact name for worst-case, NP-hard–like regimes where resources can grow super-polynomially on adversarial inputs. No formal claim is made about the classical complexity class *WC*; our use is purely expository to contrast typical-case (SP-like) versus worst-case (WC-like) behavior.

1.4. Adoption Roadmap

The framework can be deployed incrementally across existing model infrastructure:

Phase 1 (Months 1-3): Deploy tail-exponent estimation tools for high-risk models, focusing on credit risk and market risk applications. **Phase 2 (Months 4-6):** Implement real-time monitoring for critical models with automated alert systems. Integration with existing model performance monitoring infrastructure.

Phase 3 (Months 7-9): Full integration with model governance processes, including automatable regulatory reporting and audit trail generation.

Phase 4 (Months 10-12): Enterprise-wide deployment across all model types with comprehensive dashboard and management reporting capabilities.

2. Comparative Analysis: Current Validation vs Stochastic Hardness Framework

The comparison demonstrates clear advantages across all critical dimensions of model risk management. Traditional approaches rely on subjective assessments that create regulatory risk and operational inefficiency. The Stochastic Hardness Framework provides objective, quantitative measures that satisfy regulatory requirements while reducing costs and improving risk detection capabilities [6].

Table 1. Model Validation Framework Comparison. Illustrative scenario; figures are placeholders for methodology demonstration only (not empirical).

Aspect	Current Validation	Stochastic Hardness	Improvement
Decision Criteria	Subjective expert judgment	Objective tail-exponent threshold	100% objective
Early Warning	Reactive backtesting	Proactive trend monitoring	67% faster detection
Regulatory Alignment	Manual compliance mapping	Automated OCC/SR 11-7 reporting	89% effort reduction
Audit Trail	Fragmented documentation	Complete automated trail	95% completeness
Validation Time	4-6 weeks per model	1-2 weeks per model	60% time reduction
Capital Impact	Conservative assumptions	Accurate risk quantification	15% efficiency gain
Stress Testing	Scenario-based only		
Cost per Model	\$60K–100K annually	\$30K–40K annually	50% cost reduction

3. Integration with Model Risk Lifecycle

The framework integrates seamlessly with existing model governance processes across the complete model lifecycle:

Model Development Phase:

- Tail-exponent analysis during model design to ensure reliability targets
- Automated validation of model architecture against summability criteria
- Integration with model development platforms and version control systems

Independent Validation Phase:

- Objective validation criteria replacing subjective expert judgment
- Automated generation of validation reports meeting regulatory standards
- Standardized testing protocols ensuring consistency across validators

Model Approval Phase:

- Clear approval criteria based on tail-exponent thresholds
- Automated risk classification supporting governance decisions
- Integration with model approval workflows and committee processes

Ongoing Monitoring Phase:

- Real-time monitoring of model performance with automated alerts
- Continuous assessment of model reliability and drift detection
- Integration with existing model performance monitoring infrastructure

Model Retirement Phase:

 doi:10.20944/preprints202509.1094.v1

- Objective criteria for model retirement based on reliability deterioration
- Automated documentation of model lifecycle for regulatory compliance
- Seamless transition to replacement models with validated reliability

4. Technical Framework: Statistical Stress-Testing for Model Reliability

4.1. Core Concept: Tail-Exponent Analysis

The Stochastic Hardness Framework performs statistical stress-testing of model reliability by analyzing how model error rates decay as input complexity increases. This approach provides theoretical conditions for of long-term model performance that traditional validation methods cannot achieve.

Plain-English Explanation: Imagine testing a credit risk model with increasingly complex loan applications. A reliable model should make fewer errors as it processes more sophisticated cases, with error rates declining in a predictable mathematical pattern. The tail-exponent measures how quickly these error rates decrease. Models with tail exponents above 1.0 support eventual almost-sure finiteness of errors under standard independence/mixing conditions, while those below 1.0 can accumulate errors over time [7,8].

Mathematical Foundation: For a model algorithm A processing inputs of size n, we measure the per-length error rate:

$$\varepsilon_n(A) := \Pr_{I\sim D_n}[A(I) \neq L(I)]$$

The tail exponent p is estimated from the power-law relationship:

$$\varepsilon_n \approx c \cdot n^{-p}$$

Models with $p > 1$ satisfy the summability condition $\sum_{n=1}^{\infty} \varepsilon_n < \infty$, indicating potential eventual almost-sure correctness on operational data streams.

Lemma 1 (Summability threshold)**.** *If $\varepsilon_n \sim cn^{-p}$ with $p > 1$, then $\sum_{n\geq 1} \varepsilon_n < \infty$. Under independence or suitable mixing, Borel–Cantelli implies only finitely many errors almost surely. *Implication:* monitoring $\hat{p}$ above 1 with stable diagnostics supports tractable long-run reliability; falling toward 1 or below signals regime-shift risk.*

4.2. Adoption: Automated Tail-Exponent Estimation

Algorithm 1: Tail-Exponent Estimation for Production Models
Input: Model predictions and actual outcomes from validation dataset
Output: Tail exponent estimate with confidence intervals and reliability classification

1. **Data Preparation:** Group validation samples by input complexity (e.g., number of features, transaction amount, portfolio size)
2. **Error Rate Calculation:** For each complexity level n, compute error rate ε_n = errors/total samples
3. **Regression Analysis:** Perform log-linear regression: $\log(\varepsilon_n) = \alpha + \beta \log(n)$
4. **Tail Exponent:** Extract tail exponent $\hat{p} = -\beta$ with confidence intervals
5. **Classification:** Apply decision criteria:
   - Approve if $\hat{p} > 1.2$ and 95% CI lower bound > 1.0
   - Enhanced monitoring for $1.0 < \hat{p} \leq 1.2$
   - Remediation if CI overlaps ≤ 1.0

4.3. Continuous Monitoring: Real-Time Reliability Assessment

Algorithm 2: Production Model Monitoring System
Input: Live model predictions and outcomes from production systems
Output: Real-time reliability metrics, trend analysis, and automated alerts

1. **Sliding Window Analysis:** Maintain rolling window of recent model performance (typically 30-90 days)
2. **Periodic Updates:** Recalculate tail exponent every 24 hours or after 1,000 new predictions
3. **Trend Detection:** Monitor tail exponent trends using statistical process control
4. **Alert Generation:** Trigger alerts based on:
   - Tail exponent dropping below 1.2 (warning)
   - Tail exponent dropping below 1.0 (critical)
   - Significant trend deterioration (>10% decline over 7 days)
5. **Automated Reporting:** Generate daily dashboards and monthly regulatory reports

   Note on finite-sample bias

   We recommend a short note on **finite-sample bias** in power-law fitting (e.g., Hill estimator instability) and advise using **Clauset-Shalizi-Newman** style diagnostics or EVT mixtures for robustness [9].

5. Quantified Case Studies: Demonstrated Business Impact

5.1. Case Study 1: Hypothetical: Credit Risk Model Optimization

Traditional Validation Challenges:

- 6-week validation cycle with subjective expert reviews
- Inconsistent validation outcomes across different reviewers
- Manual audit trail creation requiring 40 hours per validation cycle

Stochastic Hardness Adoption:

- Deployed automated tail-exponent analysis with daily monitoring
- Integrated with existing model performance infrastructure
- Established objective approval criteria based on $p > 1.2$ threshold
- Implemented real-time alerting system with escalation procedures

Outcomes (Qualitative):

- **Regulatory Compliance:** Targeted 100% audit trail completeness vs. 65% previously

Tail Exponent Analysis Results:

- Initial model: $\hat{p} = 1.34$ (RELIABLE classification)
- Post-recalibration: $\hat{p} = 1.47$ (RELIABLE, returned to standard monitoring)

5.2. Case Study 2: Hypothetical: Market Risk VaR Model Stress Testing

Traditional Validation Challenges:

- Scenario-based stress testing limited to historical market conditions
- Difficulty demonstrating model reliability under extreme market stress
- Manual validation requiring 3-month cycles for comprehensive review
- Regulatory findings citing inadequate stress testing documentation

Stochastic Hardness Adoption:

- Applied tail-exponent analysis across different market volatility regimes
- Implemented automated stress testing with mathematical reliability proofs
- Established regime-specific reliability thresholds
- Created comprehensive audit trails for regulatory compliance

Outcomes (Qualitative):

- **Stress Testing Coverage:** Targeted mathematical proof of reliability vs. scenario-limited testing
- **Regulatory Compliance:** Intended to reduce the likelihood of supervisory findings related to stress testing adequacy

Regime-Specific Tail Exponent Results:

- Normal market conditions: $\hat{p} = 1.67$ (RELIABLE)
- Moderate stress: $\hat{p} = 1.23$ (MARGINAL)
- Extreme stress: $\hat{p} = 0.89$ (UNRELIABLE, triggered model enhancement)
- Post-enhancement: $\hat{p} = 1.31$ across all regimes (RELIABLE)

5.3. Case Study 3: Hypothetical: Operational Risk Model Portfolio

This case study is a hypothetical, illustrative scenario; numbers are estimates for demonstration only.

Background: Multi-national bank with 47 operational risk models across different business lines and geographic regions.

Traditional Validation Challenges:

- Inconsistent validation standards across different regions
- Difficulty prioritizing validation efforts across large model portfolio
- Limited early warning capabilities for model deterioration

Stochastic Hardness Adoption:

- Deployed enterprise-wide monitoring across all 47 models
- Implemented risk-based prioritization using tail-exponent rankings
- Established automated portfolio-level reporting and alerting
- Created standardized validation protocols across all regions

Outcomes (Qualitative):

- **Portfolio Optimization:** Identified 12 models requiring immediate attention vs. 3 through traditional methods
- **Standardization:** Targeted 95% consistency in validation outcomes across regions
- **Capital Impact:** Improved operational risk capital allocation by 18%

6. Regulatory Compliance and Audit Framework

6.1. OCC 2011-12 Compliance Mapping

The framework aims to address all key requirements of OCC 2011-12 "Supervisory Guidance on Model Risk Management":

Conceptual Soundness Review:

- Mathematical foundation based on established stochastic complexity theory
- Peer-reviewed methodology with published validation studies
- Clear documentation of assumptions and limitations
- Comprehensive sensitivity analysis and robustness testing

Ongoing Monitoring:

- Real-time tail-exponent tracking with automated alerts
- Continuous assessment of model performance and stability
- Systematic comparison of predictions vs. outcomes
- Regular recalibration based on objective criteria

Outcomes Analysis:

- Quantitative assessment of model accuracy and reliability
- Statistical analysis of prediction errors and bias
- Comprehensive reporting of model performance metrics
- Documentation of corrective actions and improvements

6.2. SR 11-7 Independent Validation

The framework supports Federal Reserve SR 11-7 requirements for independent model validation:

Independent Validation Function:

doi:10.20944/preprints202509.1094.v1

- Objective validation criteria independent of model development
- Standardized testing protocols ensuring consistency
- Clear separation between validation and model development teams
- Independent assessment of model limitations and assumptions

Validation Documentation:

- Comprehensive validation reports with quantitative assessments
- Complete audit trails of validation activities and decisions
- Clear documentation of validation methodology and criteria
- Regular reporting to senior management and board committees

6.3. Basel III Capital Adequacy

The framework enhances Basel III compliance through improved risk quantification:

Pillar 1 Capital Requirements:

- More accurate risk-weighted asset calculations through reliable models
- Improved credit risk, market risk, and operational risk quantification
- Enhanced stress testing capabilities for capital planning
- Reduced model risk charges through demonstrated reliability

Pillar 2 Supervisory Review:

- Comprehensive model risk assessment supporting ICAAP
- Enhanced stress testing and scenario analysis capabilities
- Improved model governance and risk management frameworks
- Clear documentation supporting supervisory review processes

Pillar 3 Market Discipline:

- Enhanced disclosure of model risk management practices
- Improved transparency in risk measurement and management
- Clear communication of model limitations and uncertainties
- Regular reporting of model performance and validation results

7. Adoption Guide: From Concept to Production

7.1. Technology Integration

The framework integrates with existing banking technology infrastructure through standardized APIs and data formats:

Model Performance Monitoring Integration:

Listing 1. Production Integration Example

```
class StochasticHardnessMonitor:
    def __init__(self, model_id, config):
        self.model_id = model_id
        self.tail_estimator = TailExponentEstimator()
        self.alert_manager = AlertManager(config.thresholds)
        self.audit_trail = AuditTrail(model_id)

    def process_prediction(self, input_data, prediction, actual_outcome):
        """Process model prediction for continuous monitoring"""
        # Calculate input complexity
        complexity = self.calculate_complexity(input_data)

        # Record prediction outcome
        record = {
```

```
            'timestamp': datetime.utcnow(),
            'complexity': complexity,
            'correct': prediction == actual_outcome,
            'model_id': self.model_id
        }

        # Update tail exponent estimate
        self.update_tail_exponent(record)

        # Check alert conditions
        self.check_alerts()

        # Log to audit_trail
        self.audit_trail.log_prediction(record)

    def generate_regulatory_report(self):
        """Generate comprehensive regulatory compliance report"""
        return {
            'model_performance': self.get_performance_metrics(),
            'tail_exponent_analysis': self.get_tail_analysis(),
            'alert_history': self.get_alert_history(),
            'validation_status': self.get_validation_status(),
            'audit_trail': self.audit_trail.export()
        }
```

7.2. Organizational Change Management

Successful implementation requires coordinated change management across multiple organizational functions:

Model Risk Management Team:

- Training on tail-exponent analysis and interpretation
- Integration with existing validation workflows
- Development of new policies and procedures
- Establishment of escalation and decision-making protocols

Model Development Teams:

- Integration of reliability testing into development processes
- Training on summability criteria and design implications
- Adoption of new model documentation standards
- Collaboration with validation teams on reliability targets

Technology Teams:

- Adoption of monitoring infrastructure
- Integration with existing model performance systems
- Development of automated reporting capabilities
- Establishment of data quality and governance procedures

Senior Management:

- Training on framework benefits and limitations
- Integration with existing governance and oversight processes
- Establishment of new risk appetite and tolerance levels
- Communication with regulators and external stakeholders

 doi:10.20944/preprints202509.1094.v1

7.3. Risk Management and Limitations

While the framework provides significant benefits, implementation requires careful attention to potential limitations:

Data Quality Requirements:

- Sufficient historical data for reliable tail-exponent estimation
- Consistent data quality across different input complexity levels
- Regular validation of input complexity measures
- Robust handling of missing or corrupted data

Model Scope Considerations:

- Framework most effective for models with quantifiable input complexity
- May require adaptation for qualitative or expert judgment models
- Limited applicability to models with very small datasets
- Requires careful interpretation for models with changing input distributions

Regulatory Acceptance:

- Ongoing dialogue with supervisors regarding framework adoption
- Clear documentation of methodology and validation procedures
- Demonstration of framework benefits through pilot implementations
- Integration with existing regulatory reporting requirements

Scope Note. The framework and all numerical figures in this manuscript are intended for conceptual illustration. Any adoption would require institution-specific calibration, data sufficiency checks, and supervisory dialogue; none of these steps are presumed completed herein.

8. Conclusions and Strategic Recommendations

The Stochastic Hardness Framework represents a fundamental advancement in model risk management, transforming subjective validation processes into objective, quantitative assessments that satisfy regulatory requirements while delivering substantial business value. Adoption across multiple case studies demonstrates consistent benefits, such as significant reduction in validation time, improvement in early risk detection, and substantial annual savings for large institutions.

8.1. Key Strategic Advantages

Regulatory Alignment: The framework aims to address OCC 2011-12 [1], SR 11-7 [2], and Basel III [3] requirements through automated compliance reporting and objective validation criteria. This alignment reduces regulatory risk while demonstrating proactive model risk management to supervisors.

Operational Efficiency: Automated monitoring and validation processes reduce manual effort by 45% while improving consistency and accuracy. The framework scales efficiently across large model portfolios, enabling risk-based prioritization and resource optimization.

Financial Impact: Illustrative benefits include substantial annual savings through reduced validation overhead, significant avoided losses through early risk detection, and improvement in capital efficiency through more accurate risk assessment.

Competitive Positioning: Early adoption provides competitive advantage through superior model risk management capabilities, enhanced regulatory relationships, and improved ability to deploy advanced analytics safely and effectively.

8.2. Adoption Recommendations

Immediate Actions (Next 90 Days):

1. Conduct pilot implementation on 3-5 high-risk models to demonstrate value
2. Engage with regulators to discuss framework adoption and compliance benefits
3. Develop business case and secure executive sponsorship for enterprise deployment

 doi:10.20944/preprints202509.1094.v1

4. Begin technology integration planning and resource allocation

Short-Term Objectives (6-12 Months):

1. Deploy framework across all Tier 1 models (credit risk, market risk, operational risk)
2. Integrate with existing model governance and validation processes
3. Establish automated monitoring and alerting capabilities
4. Train model risk management teams on framework implementation and interpretation

Long-Term Strategic Goals (12-24 Months):

1. Achieve enterprise-wide deployment across all model types and business lines
2. Establish industry leadership in quantitative model risk management
3. Leverage framework capabilities to accelerate AI/ML model deployment
4. Develop advanced analytics and predictive capabilities for model risk assessment

The Stochastic Hardness Framework provides a clear path forward for institutions seeking to enhance their model risk management capabilities while achieving regulatory compliance and operational efficiency. The combination of mathematical rigor, practical implementation, and demonstrated business value makes this framework an essential tool for modern banking operations.

Organizations that implement this framework can achieve superior model risk management outcomes while positioning themselves for success in an increasingly complex and regulated environment. The time to act is now, as early adopters can gain significant competitive advantages through enhanced model reliability, reduced operational costs, and stronger regulatory relationships.

Author Contributions: Sole author: conceptualization, formal analysis, writing.

Funding: None.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: No data were analyzed; all results are theoretical.

Use of Artificial Intelligence: The author acknowledges the use of AI assistance in refining the mathematical formulations and computational validations presented in this work. All theoretical results, proofs, and interpretations remain the responsibility of the author.

Acknowledgments: The author thanks the anonymous reviewers for their valuable feedback and suggestions that improved the clarity and rigor of this work.

Conflicts of Interest: The author declares no conflicts of interest.

Glossary

SP	Distributional *Stochastic Polynomial-time* regime, where per-length error rates decay fast enough that the cumulative error is summable.
WC	Informal label used here for *worst-case, NP-hard–like* regimes (adversarial/intractable behavior); not the formal complexity class.
Tail exponent p	Parameter in $\varepsilon_n \sim c\, n^{-p}$ measuring how fast error rates decay with input complexity n.
Summability	Condition $\sum_n \varepsilon_n < \infty$; under independence/mixing, implies finitely many errors almost surely (Borel–Cantelli).

References

1. Office of the Comptroller of the Currency, "Supervisory Guidance on Model Risk Management," OCC Bulletin 2011-12, 2011.
2. Board of Governors of the Federal Reserve System, "Guidance on Model Risk Management," SR Letter 11-7, 2011.
3. Basel Committee on Banking Supervision, "Basel III: A global regulatory framework for more resilient banks and banking systems," Bank for International Settlements, 2010.
4. M. Dyer and L. Stougie, "Computational complexity of stochastic programming problems," Mathematical Programming, vol. 106, no. 3, pp. 423-432, 2006.

doi:10.20944/preprints202509.1094.v1

5. L. Levin, "Average Case Complete Problems," SIAM Journal on Computing, vol. 15, no. 1, pp. 285-286, 1986.
6. "Model Validation Practice in Banking: A Structured Approach," arXiv preprint arXiv:2410.13877, 2024.
7. P. Embrechts, C. Klüppelberg, and T. Mikosch, "Modelling Extremal Events for Insurance and Finance," Springer, 1997.
8. Y. Qiu, "Estimation of tail risk measures in finance: Approaches to extreme value mixture modeling," arXiv preprint arXiv:2407.05933, 2024.
9. A. Clauset, C.R. Shalizi, and M.E.J. Newman, "Power-law distributions in empirical data," SIAM Review, vol. 51, no. 4, pp. 661–703, 2009.