Extreme Value Theory and Tail Risk

Application to NVDA Stock Analysis

Yinhuan Chen
McGill University

Directed Reading Program (DRP)
April 6, 2026

Abstract

Extreme Value Theory (EVT) provides a mathematical framework for modeling rare and extreme events. In this project, we study the Peaks-Over-Threshold (POT) method and apply it to NVDA stock data to detect extreme market movements and estimate tail risk. By bridging asymptotic probability theory and financial modeling, this work illustrates how mathematical theory can be applied to real-world risk analysis.

1 Introduction

Extreme events have fascinated mathematicians for over a century. While classical probability theory developed tools to understand average behavior and central tendencies, it became clear that rare and extreme observations required a different framework.

The systematic study of extremes began in the early 20th century with the work of Fisher and Tippett (1928), who characterized the limiting distributions of maxima of independent random variables. Later, Gnedenko extended these results, forming what is now known as Extreme Value Theory (EVT).

The motivation for EVT arises from real-world phenomena where rare events have disproportionately large consequences. Examples include floods, earthquakes, insurance losses, and financial market crashes. Classical Gaussian models underestimate the probability of such events due to their thin tails. Financial returns, in contrast, exhibit heavy-tailed behavior.

Extreme Value Theory provides an asymptotic framework for modeling tail behavior. Rather than describing the entire distribution, EVT focuses exclusively on extreme observations and provides universal limiting distributions under mild conditions.

In this report, I will introduce the Peaks-Over-Threshold (POT) method and apply it to NVDA stock data to analyze tail risk.

2 Theoretical Framework and Key Results

2.1 Core Idea

The central idea of the article is that extreme observations can be modeled independently of the full distribution.

Instead of modeling the entire dataset, one selects a sufficiently high threshold and studies only the exceedances above that threshold.

This approach is known as the Peaks-Over-Threshold (POT) method.

2.2 Peaks-Over-Threshold and the Generalized Pareto Distribution

Definition 1 (Exceedances)**.** *Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables with distribution function F. For a high threshold u, the exceedances are defined as*

$$Y = X - u \quad \textit{given that } X > u.$$

Theorem 1 (Pickands–Balkema–de Haan)**.** *For a sufficiently high threshold u, the conditional distribution of exceedances converges to a Generalized Pareto Distribution (GPD):*

$$G(y) = \begin{cases} 1 - \left(1 + \gamma \frac{y}{\sigma}\right)^{-1/\gamma}, & \gamma \neq 0, \\ 1 - e^{-y/\sigma}, & \gamma = 0. \end{cases}$$

2.3 Interpretation of Parameters

The GPD contains two important parameters:

- γ (shape parameter), which controls the heaviness of the tail;
- σ (scale parameter), which determines the magnitude of extreme deviations.

If $\gamma > 0$, the distribution has heavy tails and large extremes become more likely. If $\gamma = 0$, the distribution reduces to the exponential case. If $\gamma < 0$, the distribution has a finite upper bound.

2.4 Key Conceptual Points

The article emphasizes several important insights:

- EVT does not require knowledge of the full underlying distribution.
- Extreme behavior is asymptotically universal.
- Modeling the tail separately provides more accurate risk estimates.

3 Application: EVT-Based NVDA Stock Detection with Drift

3.1 Financial Motivation

Extreme Value Theory (EVT) is particularly relevant in financial risk modeling, where rare but high-impact events can significantly affect portfolios and markets. Financial return series are known to exhibit heavy tails, volatility clustering, and non-stationarity over time.

In this section, we apply a drift-adjusted Peaks Over Threshold (POT) approach to daily log-returns of NVIDIA Corporation (NVDA). The objective is to detect extreme return events while accounting for time-varying drift in the series.

3.2 Data Description

We collected daily closing prices for NVDA from January 1, 2023 to December 31, 2025 using the `quantmod` package in R.

The log-return series was computed as

$$R_t = \log(P_t) - \log(P_{t-1}),$$

where P_t denotes the closing price at time t. Log-returns are used because they stabilize variance and are standard in financial modeling.

We select the empirical 98% quantile as the threshold and model exceedances above this level.

An initialization window of 200 observations is used to estimate the GPD parameters.

3.3 Drift-Adjusted POT Methodology

Financial return series often exhibit slow time-varying trends or local mean shifts. To account for this non-stationarity, we introduce a rolling drift component.

Rolling Drift Estimation Let R_t denote the daily log-return. We estimate a time-varying drift M_t using a rolling mean:

$$M_t = \frac{1}{w} \sum_{i=t-w+1}^{t} R_i,$$

where w is the window size (here $w = 50$).

The drift-adjusted residual series is then defined as:

$$\tilde{R}_t = R_t - M_t.$$

This transformation isolates stochastic fluctuations from deterministic trend behavior.

Peaks Over Threshold (POT) We apply the POT method to the residual series $\tilde{R}_t$.

A high threshold u is selected as the 98th percentile of an initial training sample. For observations exceeding u, we define exceedances:

$$Y_i = \tilde{R}_i - u, \quad \text{for } \tilde{R}_i > u.$$

According to Extreme Value Theory, for sufficiently large u, the exceedances approximately follow a Generalized Pareto Distribution (GPD):

$$G(y) = 1 - \left(1 + \frac{\gamma y}{\sigma}\right)^{-1/\gamma},$$

where:

- γ is the shape parameter (tail index),
- σ is the scale parameter.

Dynamic Extreme Threshold After estimating $\hat{\gamma}$ and $\hat{\sigma}$, the extreme quantile z_q is computed as:

$$z_q = u + \frac{\hat{\sigma}}{\hat{\gamma}} \left[\left(\frac{N_u}{n\alpha} \right)^{\hat{\gamma}} - 1 \right],$$

where:

- N_u is the number of exceedances,
- n is the training sample size,
- α is the target tail probability (e.g., 0.01).

Finally, the dynamic anomaly detection threshold is reconstructed as:

$$T_t = M_t + z_q.$$

An observation is flagged as extreme whenever:

$$R_t > T_t.$$

3.4 R Implementation and Empirical Results

The EVT-based anomaly detection with drift adjustment was implemented in R. Daily NVDA prices were retrieved using the `quantmod` package, and log-returns were computed. A rolling mean with window size 50 was used to estimate the drift component M_t.

The POT method was then applied to the residual series $\tilde{R}_t = R_t - M_t$. The threshold was set at the 98th percentile of the first 200 observations, and exceedances were fitted using the Generalized Pareto Distribution.

R Code

```
library(quantmod)
library(evir)
library(ggplot2)
library(zoo)

getSymbols("NVDA", from="2023-01-01", to="2025-12-31")

price <- Cl(NVDA)
returns <- diff(log(price))
returns <- na.omit(returns)

window <- 50
M_t <- rollapply(returns,
                 width = window,
                 FUN = mean,
                 align = "right",
                 fill = NA)
```

```
residuals <- returns - M_t

init_size <- 200
init_data <- residuals[1:init_size]
u <- quantile(init_data, 0.98)

exceedances <- init_data[init_data > u] - u
fit <- gpd(exceedances, threshold = 0)

gamma_hat <- fit$par.ests["xi"]
sigma_hat <- fit$par.ests["beta"]

alpha <- 0.01
z_q <- u + (sigma_hat/gamma_hat) *
       ((length(exceedances)/(init_size * alpha))^gamma_hat - 1)

dynamic_threshold <- M_t + z_q

df <- data.frame(Date = index(returns),
                 Return = as.numeric(returns),
                 Drift = as.numeric(M_t),
                 Threshold = as.numeric(dynamic_threshold))

df$IsExtreme <- df$Return > df$Threshold
```

Results Interpretation The estimated Generalized Pareto Distribution (GPD) parameters were

$$\hat{\gamma} = 1.0005, \qquad \hat{\sigma} = 0.0175.$$

The shape parameter $\hat{\gamma}$ is approximately equal to 1, indicating strongly heavy-tailed behavior in NVDA log-returns. This suggests that extreme return events occur with substantially higher probability than under an exponential tail assumption. In particular, the estimated tail index implies a slow polynomial decay of extreme losses or gains, which is consistent with well-known stylized facts of financial return distributions.

The EVT-based dynamic threshold successfully identifies several extreme return observations, which correspond to periods of heightened market volatility and large price movements in NVDA. These detected anomalies align with visually prominent spikes in the return series.

By incorporating a rolling drift component M_t, the detection threshold adapts to gradual shifts in the underlying return process. This prevents false alarms caused by slow structural changes while preserving sensitivity to genuine tail events. Overall, the results demonstrate that the EVT framework provides an effective tool for modeling and detecting extreme movements in financial time series.

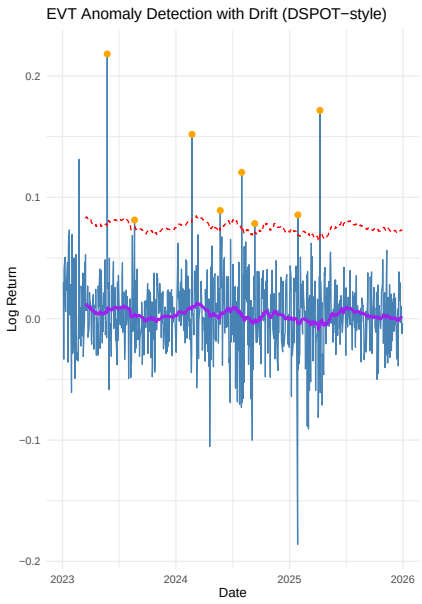


Figure 1: EVT-based anomaly detection with drift for NVDA stock returns. The blue curve represents daily log-returns, the purple curve denotes the rolling drift M_t, the red dashed line is the dynamic EVT threshold, and orange points indicate detected extreme events.

3.5 Results

The estimated shape parameter was approximately

$$\hat{\gamma} = 1.0005.$$

The positive value of $\hat{\gamma}$ indicates heavy-tailed behavior in NVDA daily log-returns. Since the estimate is close to 1, the tail decay is polynomial rather than exponential, implying a relatively high probability of extreme return events.

This suggests that large price movements occur more frequently than would be predicted under light-tailed or Gaussian assumptions. Such behavior is consistent with well-known stylized facts of financial return distributions.

Therefore, the EVT-based model captures substantial tail risk in NVDA returns at the daily frequency.

4 Conclusion

Extreme Value Theory provides a mathematically rigorous framework for analyzing rare events.

By focusing exclusively on tail behavior, EVT avoids the limitations of Gaussian models and offers improved risk estimation.

Our application to NVDA demonstrates that EVT can be used as a dynamic tail-risk detector. While the estimated tail was not strongly heavy, the method successfully identified extreme market movements.

EVT illustrates how asymptotic probability theory can be applied to real-world financial problems, bridging pure mathematics and practical risk management.