LLM Knowledge Base

Niccolò Lecce, Marco Gasparetti, Alessandro Dodon

Theoretical Foundation

Basic VaR & ES Models

Value-at-Risk at confidence level α is the threshold z_α such that with probability $1-\alpha$ the portfolio return will not fall below $-z_\alpha$. Formally:

$$\Pr(r_t < -z_\alpha) = 1 - \alpha.$$

The *historical VaR* makes no parametric assumptions: it simply takes the empirical $(1-\alpha)$-quantile of past returns. Its corresponding ES is the conditional average of losses beyond that quantile:

$$\mathrm{VaR}_\alpha^{\mathrm{hist}} = -\,\mathrm{quantile}_{1-\alpha}(r_t), \quad \mathrm{ES}_{\mathrm{hist}} = -\,\mathbb{E}[\,r_t \mid r_t < -\,z_\alpha].$$

By contrast, the *parametric (normal) VaR* assumes returns $r_t \sim N(0, \sigma^2)$:

$$\mathrm{VaR}_\alpha^{\mathrm{norm}} = z_\alpha\,\sigma, \quad \mathrm{ES}_\alpha^{\mathrm{norm}} = \sigma\,\frac{\phi(z_\alpha)}{1-\alpha},$$

with ϕ the standard normal density. For a multi-asset portfolio with monetary positions x_t and covariance matrix Σ, the portfolio VaR/ES scale in the same way:

$$\mathrm{VaR}_t = z_\alpha\,\sqrt{x_t^\top \Sigma\, x_t}, \quad \mathrm{ES}_t = \frac{\mathrm{VaR}_t}{1-\alpha}\,\phi(z_\alpha).$$

Pros

- **Historical**: fully non-parametric, captures empirical distribution and fat tails.
- **Normal**: closed-form, instant computation, easy to implement and explain.
- **Scalability**: both methods scale simply with holding period ($\sqrt{h}$) and portfolio value.

Cons

- **Historical**: may overfit past regimes, poor forward-looking properties, sensitive to sample selection.

- **Normal**: underestimates extreme losses when returns deviate from Gaussian, ignores skewness and kurtosis.

- **Static**: no time-varying volatility or correlation, cannot capture clustering or regime shifts.

Extreme Value Theory (EVT)

Extreme Value Theory targets the tail directly. Using the Peaks-Over-Threshold (POT) approach, one selects a high threshold u (e.g. the 99th-percentile of losses) and models the excesses $r_t - u$ above that threshold by a Generalized Pareto Distribution (GPD). The tail quantile and expected shortfall then admit analytic forms:

$$\widehat{\mathrm{VaR}}_\alpha = u + \frac{\hat{\beta}}{\hat{\xi}}\left[\left(\tfrac{n}{n_u}(1-\alpha)\right)^{-\hat{\xi}} - 1\right], \quad \widehat{\mathrm{ES}}_\alpha = \frac{\widehat{\mathrm{VaR}}_\alpha + \hat{\beta} - \hat{\xi}\,u}{1-\hat{\xi}},$$

where $\hat{\xi}, \hat{\beta}$ are the GPD shape and scale estimates, n the total sample size, and n_u the number of exceedances.

Pros

- Provides theoretically justified tail estimates beyond empirical maxima.

- Adapts naturally to focus on high-confidence levels (e.g. 99.9%).

- Captures heavy-tail behavior explicitly, reducing underestimation of extreme losses.

Cons

- Highly sensitive to threshold choice u, requiring diagnostic plots and stability checks.

- Needs sufficient tail data—few exceedances lead to large estimation variance.

- More complex implementation and parameter tuning than simple VaR.

Volatility Models

Modeling time-varying volatility is crucial to capture clustering and conditional risk.

GARCH(1,1)

$$\sigma_t^2 = \omega + \alpha\, r_{t-1}^2 + \beta\, \sigma_{t-1}^2, \quad r_t = \sigma_t\, z_t,$$
$$\mathrm{VaR}_{t,\alpha} = -\,\hat{\sigma}_t\, z_\alpha.$$

GARCH forecasts allow multi-period VaR via analytic recursions:

$$\mathbb{E}[\sigma_{t+\tau}^2] = \mathrm{VL} + (\alpha + \beta)^\tau (\sigma_t^2 - \mathrm{VL}),$$

where $\mathrm{VL} = \omega/(1 - \alpha - \beta)$ is the long-run variance.

Pros

- Dynamically captures volatility clustering and mean reversion.
- Enables time-consistent VaR forecasts and can be estimated efficiently via MLE.
- Forms the basis for multivariate and advanced volatility extensions.

Cons

- Assumes symmetric shocks; does not capture leverage unless extended.
- Model risk if parameters drift or structural breaks occur.
- Requires careful diagnostic checks for ARCH effects and residuals.

Extensions (EGARCH, GJR-GARCH, APARCH, ARCH(p), MA, EWMA) These variations introduce asymmetry (EGARCH, GJR), power terms (APARCH), or simple rolling/decay weights (MA, EWMA).

Pros

- Handle asymmetry, leverage effects, heavy tails.
- EWMA/MA are lightweight, requiring minimal data.

Cons

- More parameters increase overfitting risk.
- Simple MA/EWMA ignore cross-asset dependencies.

Time-Varying Correlation Models

Beyond volatility, dynamic correlations matter for portfolio aggregation. With a time-varying covariance matrix Σ_t (e.g. DCC-GARCH omitted here for simplicity):

$$\mathrm{VaR}_t = z_\alpha \sqrt{x_t^\top \Sigma_t \, x_t}.$$

Pros

- Reflects changing relationships between assets under stress.

Cons

- Significantly higher computational cost.
- Normal-joint assumption may still underestimate tail dependence.

Factor Models

Factor models reduce dimensionality by attributing risk to common drivers.

Single-Factor (Sharpe) Model

$$r_i = \alpha_i + \beta_i \, r_m + \varepsilon_i, \quad \sigma_p^2 = \Big(\sum_i w_i \beta_i\Big)^2 \sigma_m^2 + \sum_i w_i^2 \, \sigma_{\varepsilon_i}^2,$$

$$\mathrm{VaR}_{t,\alpha} = z_\alpha \, \sigma_p, \quad \mathrm{ES}_{t,\alpha} = \sigma_p \, \frac{\phi(z_\alpha)}{1-\alpha}.$$

Pros

- Intuitive decomposition into market and idiosyncratic risk.
- Reduces dimensionality for large portfolios.

Cons

- Omits other systematic factors.
- May misestimate risk if factor model mis-specifies exposures.

Fama–French Three-Factor

$$r_i - r_f = \alpha_i + \beta_{i,1}\text{Mkt-RF} + \beta_{i,2}\text{SMB} + \beta_{i,3}\text{HML} + \varepsilon_i, \quad \sigma_p^2 = w^\top (B \Sigma_f B^\top + \Sigma_\varepsilon) \, w.$$

Pros

- Incorporates size and value factors, improving explanatory power.

Cons

- Requires robust estimation of multiple factor returns.

- Still may omit macro or sector-specific risks.

Non-Parametric Simulation

Historical & Bootstrap Simulate P&L by re-applying empirical return vectors:

$$S^{(i)} = S_0 \circ (1 + r_{\pi(i)}), \quad \Delta P^{(i)} = w^\top (S^{(i)} - S_0).$$

Compute

$$\mathrm{VaR}_\alpha = -\operatorname{quantile}_\alpha(\Delta P), \quad \mathrm{ES}_\alpha = -\mathbb{E}[\Delta P \mid \Delta P \le -\mathrm{VaR}_\alpha].$$

Pros

- No distributional assumptions, captures empirical dependence.

- Bootstrapping enhances robustness with multiple resamples.

Cons

- Computationally heavy for large scenario counts.

- Historical scenario bias, limited coverage of future states.

Backtesting

Backtesting evaluates model accuracy by comparing observed violations N over T days to expected rates.

Kupiec (Unconditional Coverage)

$$\mathrm{LR}_{\mathrm{uc}} = -2\Big[\ln((1-p)^{T-N}p^N) - \ln((1-\hat{p})^{T-N}\hat{p}^N)\Big], \quad \hat{p} = \tfrac{N}{T}.$$

Christoffersen (Conditional Coverage)

$$\mathrm{LR}_c = -2\big(\ln L_0 - \ln L_1\big).$$

Joint Test

$$\mathrm{LR}_{\mathrm{joint}} = \mathrm{LR}_{\mathrm{uc}} + \mathrm{LR}_{c}.$$

Interpretation of p-Values

- **Kupiec p-value**: if $> \alpha$, exception rate is consistent with model.
- **Christoffersen p-value**: if $> \alpha$, exceptions are serially independent.
- **Joint p-value**: if $> \alpha$, model passes both coverage and independence tests.

Violation Metrics

$$\text{Number of Violations } N, \quad \text{Violation Rate } = \frac{N}{T}.$$