Forecasting Metrics: Formal Definitions

1 Bias (Systematic Over/Under-Prediction)

Definition: Bias measures the systematic tendency of a model to over- or under-predict the actual values.

$$\text{Bias} = \frac{1}{n} \sum_{t=1}^{n} (\hat{y}_t - y_t)$$

Interpretation:

- Positive bias $\rightarrow$ model overestimates
- Negative bias $\rightarrow$ model underestimates
- Bias $\approx 0 \rightarrow$ model is well-centered

2 Calibration Error (via CRPS)

Definition: The Continuous Ranked Probability Score (CRPS) measures the quality of probabilistic forecasts by comparing the predicted CDF $F(z)$ against the actual outcome y_t:

$$\text{CRPS}(F, y_t) = \int_{-\infty}^{\infty} \left(F(z) - \mathbf{1}(z \geq y_t) \right)^2 dz$$

Interpretation:

- Captures both sharpness and calibration
- Lower CRPS = better forecast distribution

3 Residual Anomaly Detection

Definition: An anomaly is identified if the forecast error (residual) exceeds a predefined threshold:

$$r_t = y_t - \hat{y}_t, \quad \text{Anomaly if } |r_t| > \tau$$

Interpretation:

- Highlights model failures when actual deviates significantly from predicted
- Useful for real-time error monitoring

4 Data Anomaly Detection

Definition: Anomaly points in actual observed data are identified based on statistical deviation from local or global structure (e.g., z-score or IQR rule).

Interpretation:

- Independent of model output
- Detects natural spikes or drops in input data

5 Data Drift Score (Custom Distribution-Based)

Definition: A drift score quantifies deviation between reference and observed distributions via stepwise penalties:

$$\text{Drift Score} = \sum_{i=1}^{n} \text{penalty}(O_i - R_i)$$

Interpretation:

- Acceptable Range: $|O_i - R_i| < \epsilon_1$ (no penalty)
- Step Error: $\epsilon_1 \leq |O_i - R_i| < \epsilon_2$
- Penalty Error: $|O_i - R_i| \geq \epsilon_2$

6 Turning Point F1 Score (Trend Change Detection)

Definition: F1 score is adapted to evaluate detection of local minima or maxima (turning points):

$$\text{Precision} = \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN}, \quad F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Interpretation:

- Measures timing accuracy of trend reversals
- Valuable in finance, demand forecasting, economics