* **分布特征（统计学）：** 这类特征从数据整体的统计分布角度来描述波形。
* **均值 (Mean):**
* **代表什么：** 代表波形数据的平均值，也称为直流分量。
* **为何选取：** 在原始数据中，均值可以反映波形是否存在直流偏置。但在我们的预处理步骤中，已经通过“基线校正”将所有波形的均值调整至接近0。因此，对于**处理后**的数据，该特征的区分能力会大大降低，但它可以作为检验我们预处理步骤是否成功的一个指标。
* **方差 (Variance):**
* **代表什么：** 衡量数据点围绕其均值的离散程度，反映了波形交流分量的能量或波动幅度。
* **为何选取：** 在我们将波形幅度归一化到 [-1, 1] 区间后，方差成为了一个衡量波形“饱满度”的指标。例如，梯形波因为有更多的数据点分布在峰值（-1和1）附近，其方差通常会比主要数据点集中在0附近且平滑过渡的正弦波要大。因此，方差可以为区分不同波形提供有用的信息。
* **峰度 (Kurtosis):**
* **代表什么：** 衡量数据分布的尖锐或平坦程度。正态分布的峰度为0。
* **为何选取：** 三角波波形尖锐，其数据点更多地分布在峰值和谷值附近，因此呈现出“尖顶厚尾”的分布，具有很高的正峰度。梯形波因为有“平台期”，大量数据点集中在某个固定值上，分布形态比正态分布更“平坦”，呈现负峰度。正弦波则介于两者之间。因此，峰度是区分这三者的一个强力特征。
* **偏度 (Skewness):**
* **代表什么：** 衡量数据分布的不对称性。
* **为何选取：** 理想情况下三种波形都是对称的，偏度应接近0。计算该特征主要是为了检验实际数据的质量，排除因采集问题导致的波形偏移，增强模型的鲁棒性。
* **波峰因子 (Crest Factor):**
* **代表什么：** 波形峰值（绝对值最大值）与其有效值（RMS）的比率。
* **为何选取：** 这是一个直接反映波形形状的经典指标。不同形状的波形具有理论上不同的波峰因子（理想正弦波为$\sqrt{2}\approx1.414$，理想三角波为$\sqrt{3}\approx1.732$）。这个特征为模型提供了一个非常明确的几何判据。
* **最大值、最小值、峰-峰值（Peak-to-Peak, Bpp）**
* **代表什么：**幅值区间。
* **为何选取：**归一化后主要用于描述波形强度差异，是判断数据异常的重要指标。
* **形状特征（时域）：** 这类特征直接从波形随时间变化的形态入手。
* **基于导数的特征：** 我们通过计算相邻采样点的一阶差分来近似波形的导数（即变化率）。
* **差分统计量：**均值、标准差、极值、分位数。
* **为何选取：**正弦波 dB/dt 呈连续变化；三角波几乎恒定两段（正/负常值）；梯形含大段零导数。
* **导数值种类：**
* **代表什么：** 计算导数序列中不同数值的个数。
* **为何选取：** 这是识别**三角波**的**核心特征**。理想三角波的上升段和下降段斜率恒定，其导数理论上只在两个值（一个正数，一个负数）之间切换。而正弦波的导数（余弦波）是连续变化的，梯形波的导数有三个主要值（正斜率、负斜率、0）。因此，三角波的导数值种类最少。
* **二阶差分（拐点密度）**
* **代表什么：**统计二阶差分的尖峰数量和分布。
* **为何选取**：三角和梯形在拐点处二阶差分集中；正弦平滑拐点少。
* **线性段覆盖率**：
* **代表什么：**通过滑动窗口线性拟合 𝑅2判定直线段比例。
* **为何选取：**三角波几乎全程线性，梯形为“线性+平台”，正弦线性段比例最小。
* **频域特征 (备选):**
* **谐波失真度 (THD):**
* **代表什么：** 衡量一个波形相对于理想正弦波的畸变程度。
* **为何选取：** 理想正弦波不含谐波，其THD为0。而三角波和梯形波都由基波和一系列奇次谐波构成。单一指标即可区分正弦（THD≈0）、三角（中等）、梯形（较大）。因此，该特征能非常有效地将**正弦波**从另外两种波形中区分出来。
* **谐波衰减斜率**：
* **代表什么：**对 log\_An和 log\_n 线性回归的斜率，它量化了信号能量在频域上的分布趋势
* **为何选取：**三角波谱斜率≈-2，梯形波≈-1，正弦极陡趋近 -∞
* **谱熵 / 谱平坦度（Spectral Entropy / Flatness）**
* **代表什么：**频谱能量分布均匀性的度量。
* **为何选取：**正弦：能量集中，熵低；三角：较平缓；梯形：能量扩展最广，熵高。