

时间序列ARIMA

- **自回归模型（AR模型）**：用同一变量例如 x 的之前各期，即 x_1 至 x_{t-1} 来预测本期 x_t 的表现，并假设它们之间是线性关系；这是从线性回归发展而来的，只是不用 x 预测 y ，而是**用 x 预测 x （自己）**，所以叫自回归

$$\text{AR}(p)\text{定义: } X_t = c + \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t, \text{ 其中}$$

- c 是常数项
 - ε_t 被假设为均值等于0，方差等于 σ^2 的随机误差项； σ^2 被假设为对任何 t 都不变
 - 文字描述： X 的当前值等于一个或数个历史期的线性组合，加常数项，加随机误差项
- **移动平均模型（MA模型）**：自回归模型描述的是当前值与历史值之间的关系，而移动平均模型描述的是**自回归部分的误差累计**；与AR模型不同，MA模型**总是平稳的**

$$\text{MA}(q)\text{定义: } X_t = \mu + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i}, \text{ 其中}$$

- μ 是序列均值
 - $\theta_1, \dots, \theta_q$ 是参数
 - $\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q}$ 都是白噪声
- **自回归移动平均模型（ARMA模型）**： $X_t = c + \varepsilon_t + \sum_{i=1}^p \phi_i X_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j}$
- **差分自回归移动平均模型（ARIMA模型）**：ARIMA(p, d, q)中，AR是“自回归”， p 为自回归项数；MA是“移动平均”， q 为移动平均项数； d 是**使之成为平稳序列所做的差分次数（阶数）**， d 阶差分就是**相距 d 期的两个序列值相减**
- **平稳序列**：围绕一个常数上下波动且波动范围有限，即有常数均值和常数方差，检验方法有三种：

- 一是根据时序图，如果有明显的趋势或周期性，通常不是一个平稳序列
 - 二是根据自相关图和偏自相关图，平稳序列通常要么拖尾，要么截尾
 - 三是单位根检验方法，检验序列中是否存在单位根，如果存在就是非平稳序列
- **ARIMA建模流程**

1. 根据时序图、自相关函数和偏自相关函数图识别其平稳性
2. 对非平稳时间序列进行平稳化处理
3. 根据所识别出来的特征建立相关的时间序列模型（AR：ACF拖尾PACF截尾；MA：ACF截尾PACF拖尾；ARMA：ACF和PACF均拖尾）
4. 参数估计，检验是否具有统计意义
5. 假设检验，判断残差序列是否为白噪声序列
6. 利用已通过检验的模型进行预测