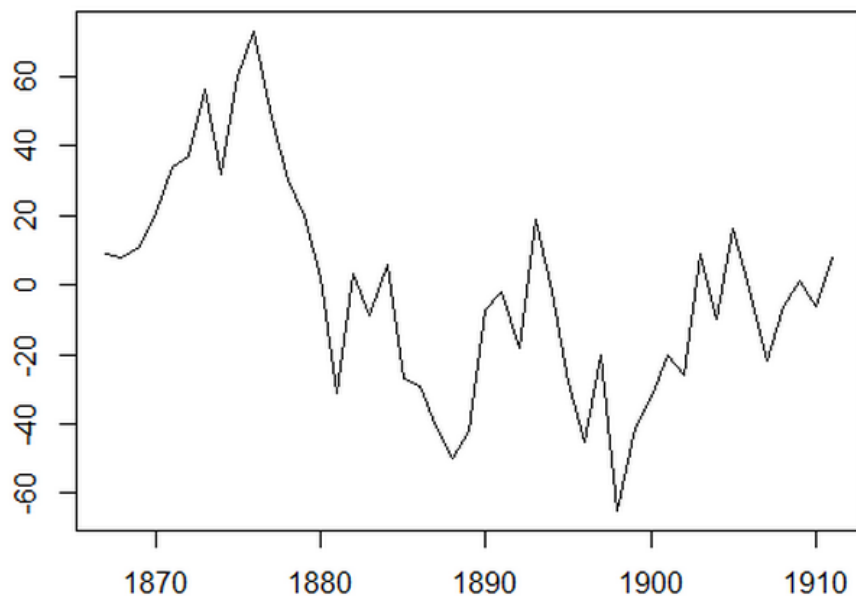


# ARIMA

✓ 平稳性：

✎ 平稳性就是要求经由样本时间序列所得到的拟合曲线在未来的一段期间内仍能顺着现有的形态“惯性”地延续下去

✎ 平稳性要求序列的均值和方差不发生明显变化



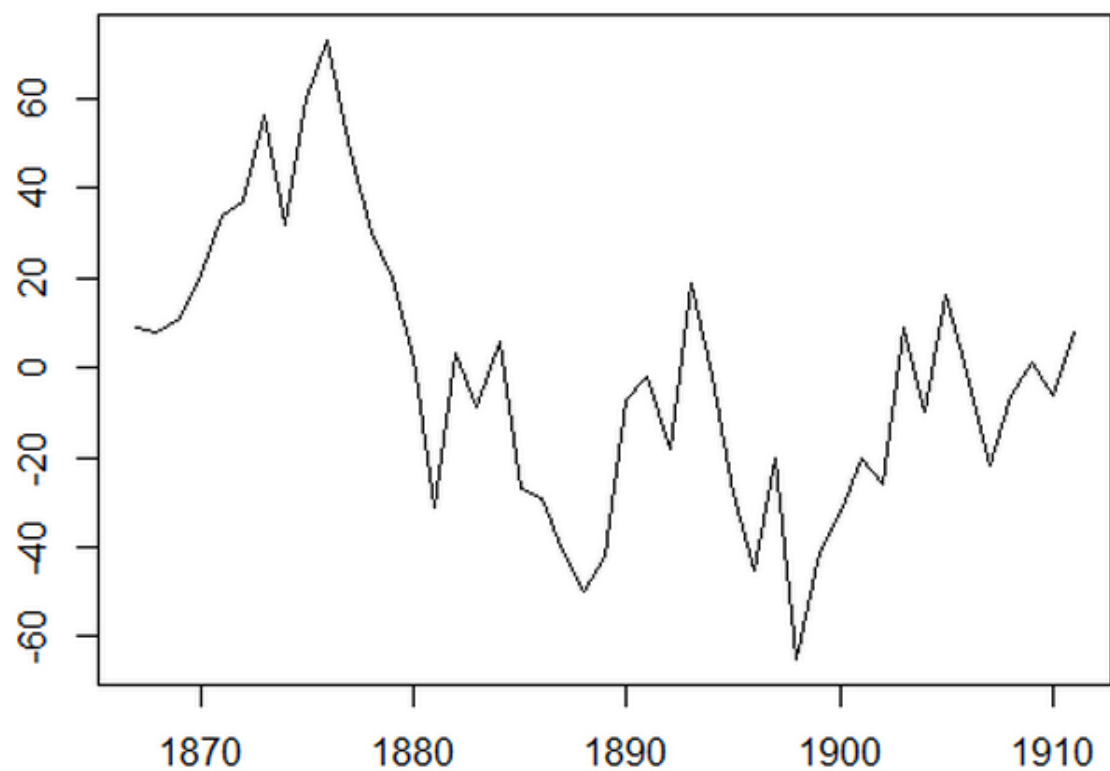
# ARIMA

## ✓ 严平稳与弱平稳：

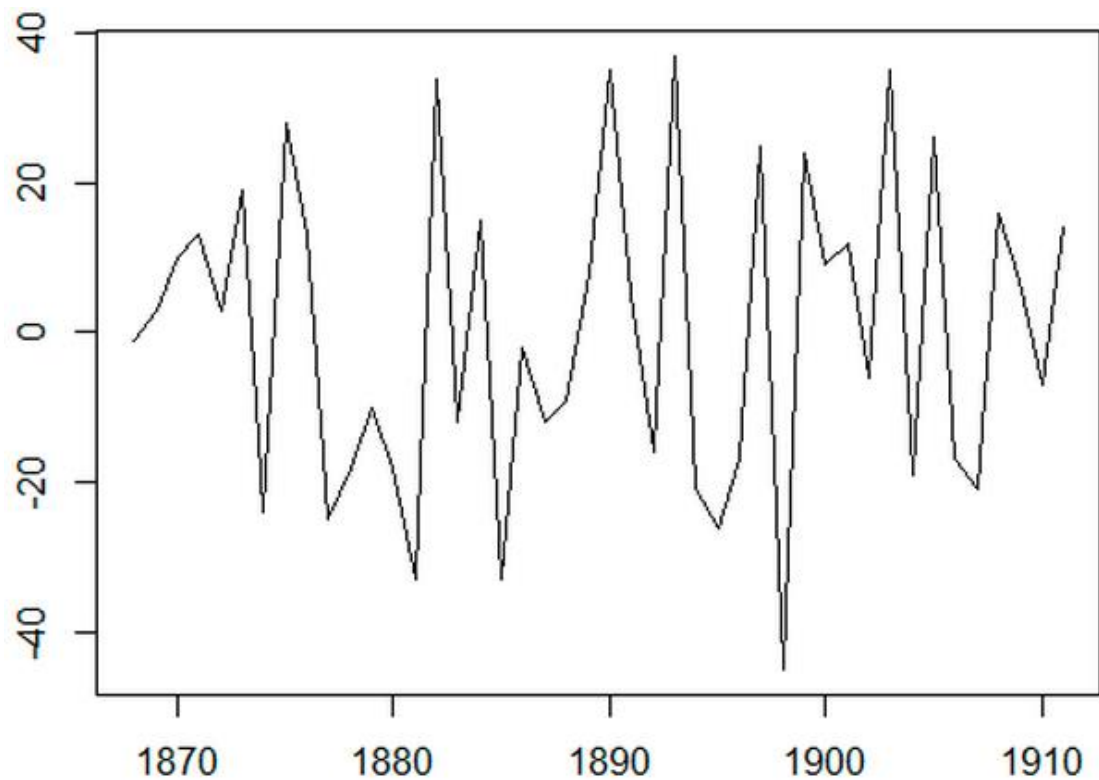
- ✎ 严平稳：严平稳表示的分布不随时间的改变而改变。  
如：白噪声（正态），无论怎么取，都是期望为0，方差为1
- ✎ 弱平稳：期望与相关系数（依赖性）不变  
未来某时刻的t的值 $X_t$ 就要依赖于它的过去信息，所以需要依赖性

# ARIMA

✓ 差分法：时间序列在 $t$ 与 $t-1$ 时刻的差值



原数据



一阶差分

# ARIMA

## ✓ 自回归模型 ( AR )

✎ 描述当前值与历史值之间的关系，用变量自身的历史时间数据对自身进行预测

✎ 自回归模型必须满足平稳性的要求

✎ p阶自回归过程的公式定义： $y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \epsilon_t$

✎  $y_t$  是当前值  $\mu$  是常数项 P 是阶数  $\gamma_i$  是自相关系数  $\epsilon_t$  是误差

# ARIMA

---

- ✓ 自回归模型的限制
  - ✎ 自回归模型是用自身的数据来进行预测
  - ✎ 必须具有平稳性
  - ✎ 必须具有自相关性，如果自相关系数( $\phi_i$ )小于0.5，则不宜采用
  - ✎ 自回归只适用于预测与自身前期相关的现象

# ARIMA

## ✓ 移动平均模型 ( MA )

✎ 移动平均模型关注的是自回归模型中的误差项的累加

✎ q阶自回归过程的公式定义： $y_t = \mu + \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i}$

✎ 移动平均法能有效地消除预测中的随机波动

# ARIMA

✓ 自回归移动平均模型 ( ARMA )

✎ 自回归与移动平均的结合

✎ 公式定义：
$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i}$$

# ARIMA

- ✓ ARIMA( $p, d, q$ )模型全称为差分自回归移动平均模型 (Autoregressive Integrated Moving Average Model, 简记ARIMA)
- ✎ AR是自回归,  $p$ 为自回归项; MA为移动平均  $q$ 为移动平均项数,  $d$ 为时间序列成为平稳时所做的差分次数
- ✎ 原理: 将非平稳时间序列转化为平稳时间序列然后将因变量仅对它的滞后值以及随机误差项的现值和滞后值进行回归所建立的模型



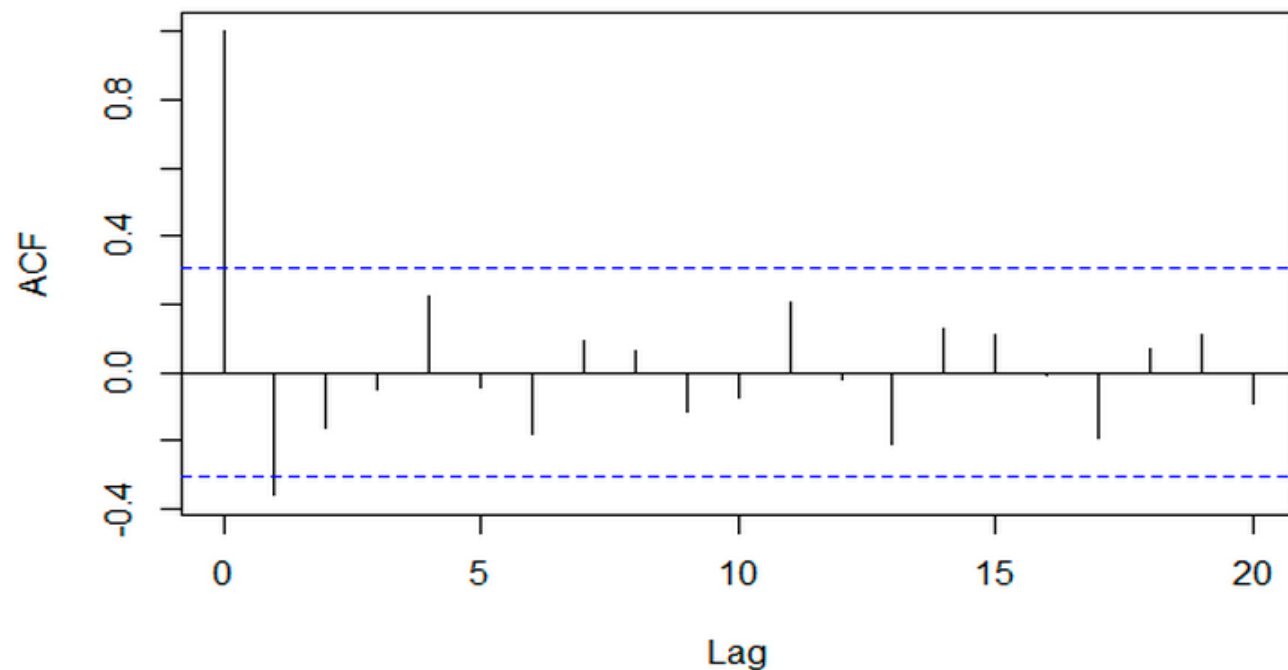
# ARIMA

✓ 自相关函数ACF(autocorrelation function)

✎ 有序的随机变量序列与其自身相比较  
自相关函数反映了同一序列在不同时序的取值之间的相关性

✎ 公式： $ACF(k) = \rho_k = \frac{\text{Cov}(y_t, y_{t-k})}{\text{Var}(y_t)}$

✎  $\rho_k$ 的取值范围为 $[-1, 1]$



- ✓ 偏自相关函数(PACF)(partial autocorrelation function)
- ✎ 对于一个平稳AR(p)模型，求出滞后k自相关系数 $p(k)$ 时实际上得到并不是 $x(t)$ 与 $x(t-k)$ 之间单纯的相关关系
- ✎  $x(t)$ 同时还会受到中间 $k-1$ 个随机变量 $x(t-1)$ 、 $x(t-2)$ 、.....、 $x(t-k+1)$ 的影响而这 $k-1$ 个随机变量又都和 $x(t-k)$ 具有相关关系所以自相关系数 $p(k)$ 里实际掺杂了其他变量对 $x(t)$ 与 $x(t-k)$ 的影响
- ✎ 剔除了中间 $k-1$ 个随机变量 $x(t-1)$ 、 $x(t-2)$ 、.....、 $x(t-k+1)$ 的干扰之后 $x(t-k)$ 对 $x(t)$ 影响的相关程度。
- ✎ ACF还包含了其他变量的影响而偏自相关系数PACF是严格这两个变量之间的相关性

# ARIMA

✓ ARIMA(p , d , q)阶数确定：

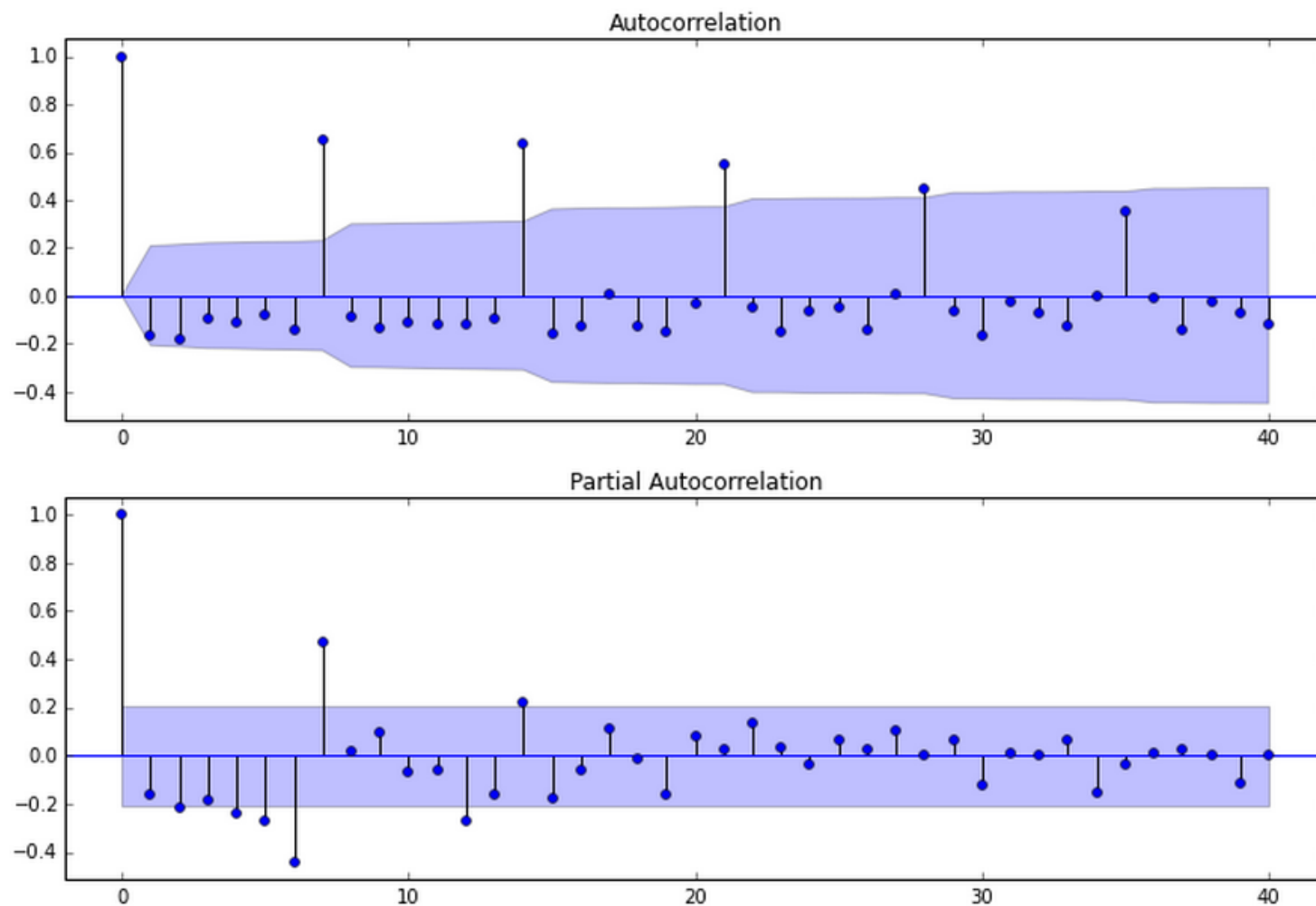
模型	ACF	PACF
AR(p)	衰减趋于零（几何型或振荡型）	p阶后截尾
MA(q)	q阶后截尾	衰减趋于零（几何型或振荡型）
ARM A(p , q)	q阶后衰减趋于零（几何型或振荡型）	p阶后衰减趋于零（几何型或振荡型）

✎ 截尾：落在置信区间内（95%的点都符合该规则）

# ARIMA

✓ ARIMA(p, d, q)阶数确定：

✎ AR(p) 看PACF  
MA(q) 看ACF



# ARIMA

---

✓ ARIMA建模流程：

✎ 将序列平稳（差分法确定d）

✎ p和q阶数确定：ACF与PACF

✎ ARIMA ( p,d,q )

# ARIMA

✓ 模型选择AIC与BIC: 选择更简单的模型

✎ AIC:赤池信息准则 ( Akaike Information Criterion , AIC )

$$AIC = 2k - 2\ln(L)$$

✎ BIC:贝叶斯信息准则 ( Bayesian Information Criterion , BIC )

$$BIC = k\ln(n) - 2\ln(L)$$

✎ k为模型参数个数，n为样本数量，L为似然函数

# ARIMA

---

✓ 模型残差检验：

✎ ARIMA模型的残差是否是平均值为0且方差为常数的正态分布

✎ QQ图：线性即正态分布