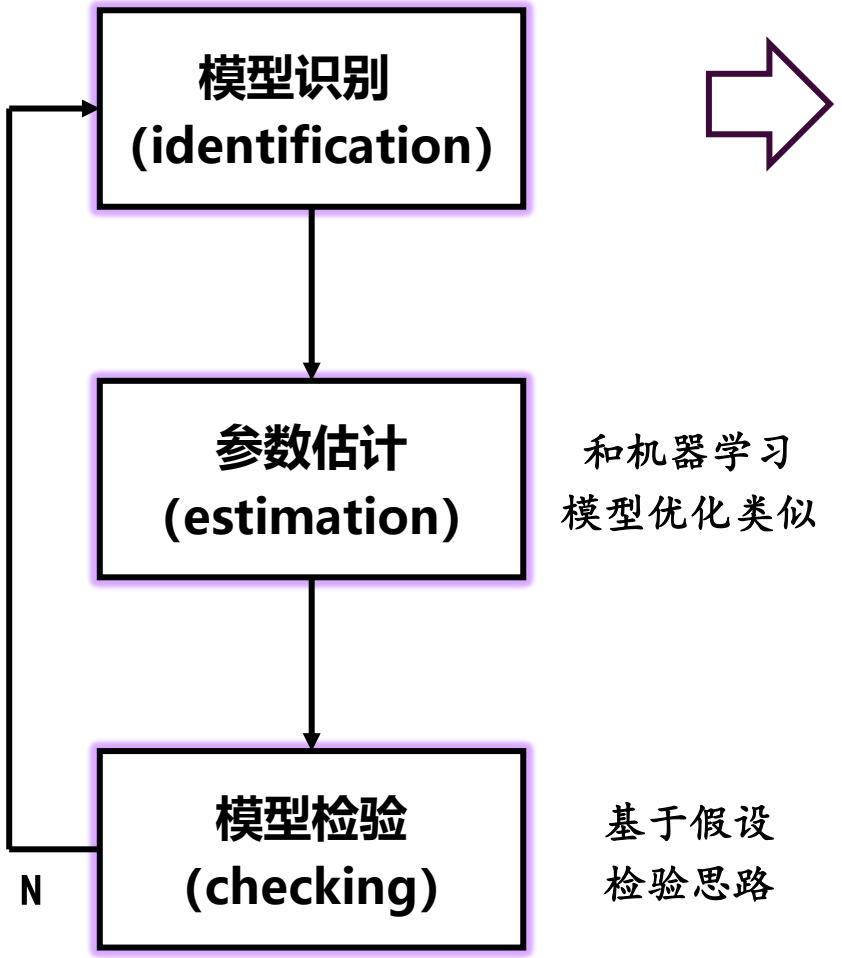
统计时间序列方法



和机器学习
模型优化类似

基于假设
检验思路

时间序列的
统计假设

面向时间序列
的统计量

面向时间序列
的概率模型

线性过程
AR, MA
ARMA,
ARIMA
…

多元时间序列模型

季节因素分解

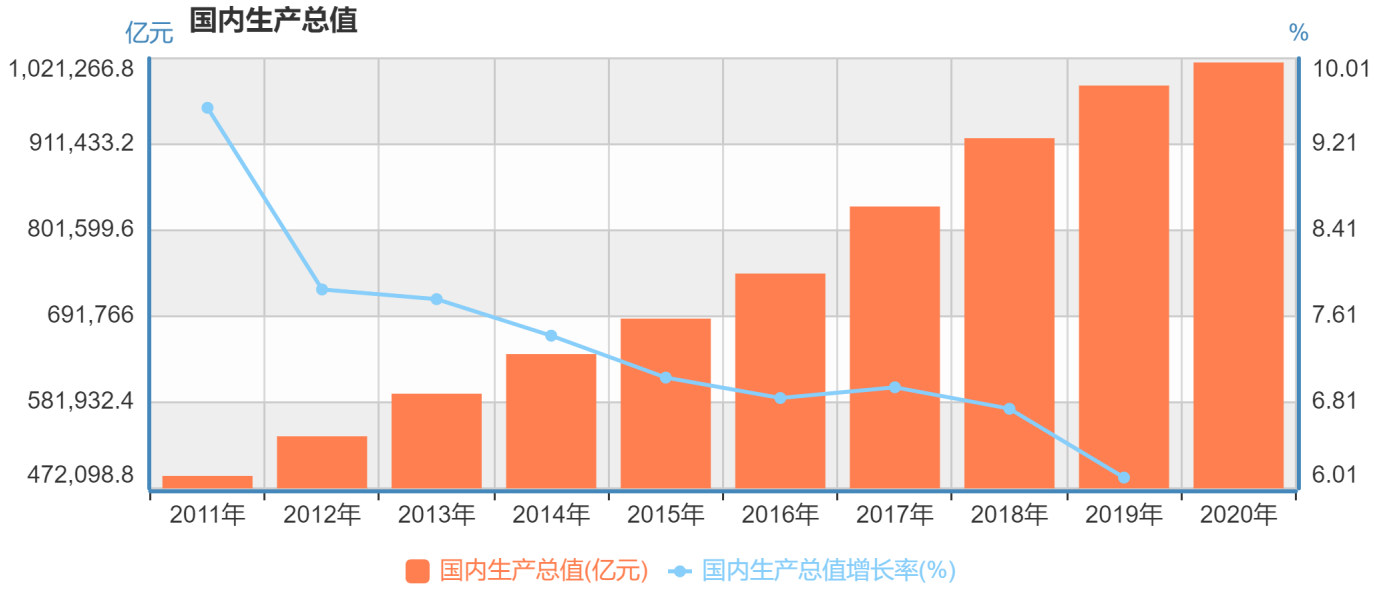
…

NJUAI 南京大学 人工智能学院
SCHOOL OF ARTIFICIAL INTELLIGENCE, NANJING UNIVERSITY
南京大學 NANJING UNIVERSITY 1902

04 时间序列的随机过程视角

时间序列的“平稳”性

时间序列的概率化定义



- 观察值序列：序列长度为n的观察值序列（一般以时间为单位）

$$y_1, y_2, \cdots, y_n$$

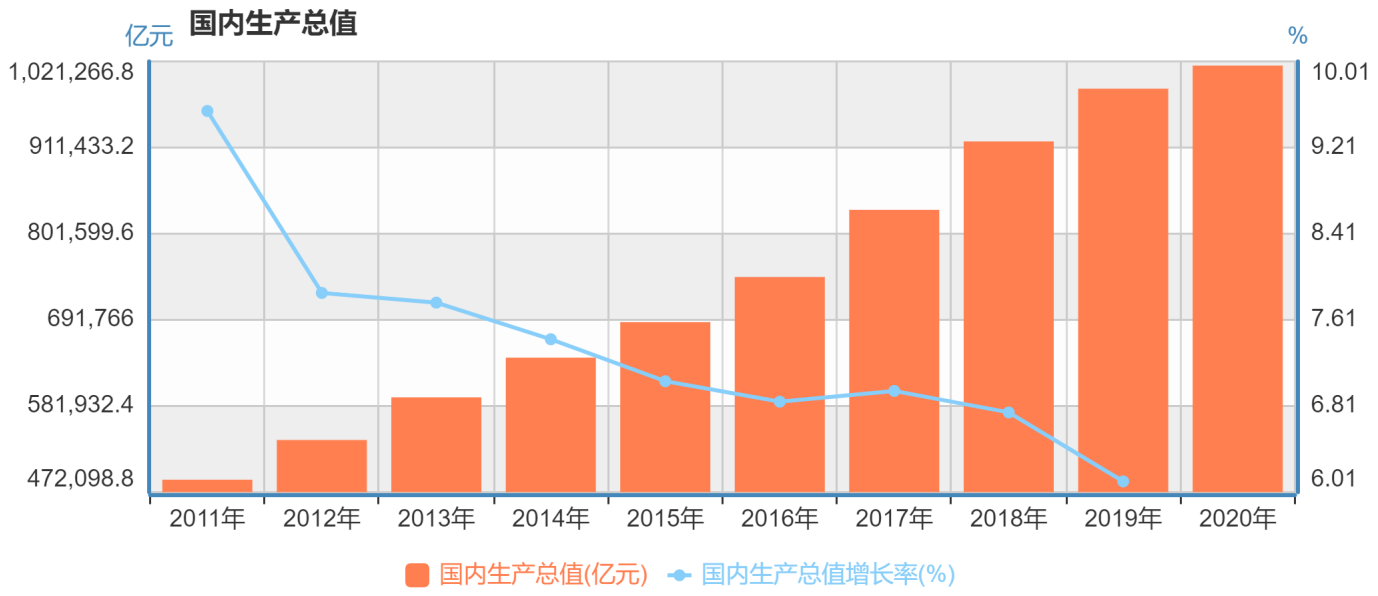
- 随机序列: 按时间顺序排列的一组**随机变量**

$$\cdots, Y_1, Y_2, \cdots, Y_t, \cdots$$

- 随机序列和观察值序列的关系
  - 观察值序列是随机序列的一个实现
  - 目的是想揭示随机时序的性质
  - 通过分析观察值序列的性质，由观察值序列的性质来推断随机序列的性质

和一般概率分布的区别是什么？

分析的基本思路

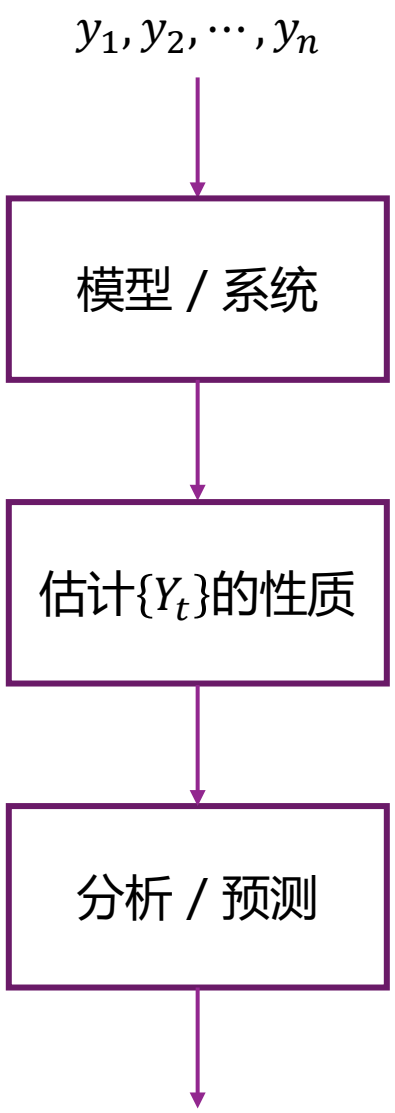


随机序列: 按时间顺序排列的一组**随机变量**

$$\cdots, Y_1, Y_2, \cdots, Y_t, \cdots$$

观察值序列：序列长度为n的观察值序列（一般以时间为单位）

$$y_1, y_2, \cdots, y_n$$



直接建模概率分布的难点

- 随机变量族的统计特性完全由它们的**联合分布**函数族$f_{Y(t_1),\dots,Y(t_n)}(y_1,\dots,y_n), n=1,2,\dots$决定，需要通过有限分布估计概率分布族的性质
- 传统统计分析：有限个变量，每个变量有多个观察值

	Y_1	…	Y_m
1	y_{11}	…	y_{m1}
2	y_{12}	…	y_{m2}
…	…	…	…
n	y_{1n}	…	y_{mn}

估计每个时刻的概率分布（或统计量）非常困难

- 时间序列：可列多个随机变量（对应不同时刻），而每个变量只有一个样本观察

	Y_1	…	Y_t
1	y_1	…	y_t

概要

1. 平稳序列的定义：

使基于有限数据的建模可行

2. 时间序列的统计量：

通过哪些统计量刻画时间序列？

3. 时间序列分析化简：

时间序列初步建模思路

分析“平稳”的时间序列

随机过程涉及的**所有**有限维联合概率分布不随时间变化

- **严平稳（Strongly Stationary）**
  - $\forall n, \forall(t_1, t_2, \dots, t_n), \forall h$，有

$$f_{Y(t_1),\cdots,Y(t_n)}(y_1, \cdots, y_n) = f_{Y(t_1+h),\cdots,Y(t_n+h)}(y_1, \cdots, y_n)$$

  - 严平稳是一种条件**比较苛刻**的平稳性定义，它认为只有当序列**所有的统计性质**都不会随着时间的推移而发生变化时，该序列才能被认为平稳。

  - 实际中大多难以实现严平稳的要求，因此需要放松对不同时刻随机变量关系的约束

如何描述一个随机过程

- 传统的统计量
  - 均值

$$\mu = E[Y] = \int_{-\infty}^{\infty} y\, f(y) dy$$

  - 协方差

$$\mathrm{Cov}[Y, X]$$
$$= E[(Y - \mu_Y)(X - \mu_X)]$$

  - 方差

$$V[Y] = E(Y - \mu)^2$$

- 时间序列的统计量
  - 均值（序列）

$$\mu(t) = E[Y(t)] = \int_{-\infty}^{\infty} y f_{Y(t)}(y) dy$$

  - **自协方差**（Autocovariance）

$$\gamma_{YY}(t_1, t_2) = \gamma(t_1, t_2) = \mathrm{Cov}[Y(t_1), Y(t_2)]$$
$$= E\big[\big(Y(t_1) - \mu(t_1)\big)\big(Y(t_2) - \mu(t_2)\big)\big]$$

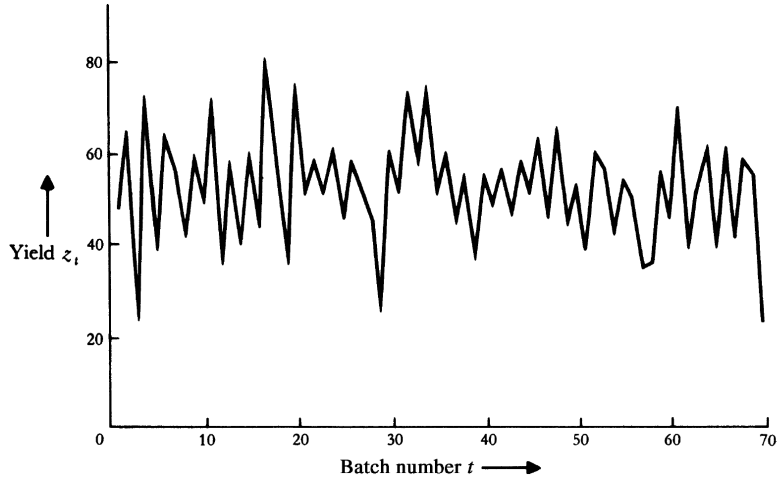
  - 方差（序列）

$$\sigma^2(t) = \gamma_{YY}(t, t)$$
$$= V[Y(t)] = E[(Y(t) - \mu(t))^2]$$

自协方差系数以及自相关系数

- 在严平稳条件下，分布$f_{Y(t_1),Y(t_2)}(y_1,y_2)$对于不同的$t_1,t_2$保持不变，因此，对具有$k$个时间间隔（$lag\ k$）的$Y_t,Y_{t+k}$，其协方差保持不变

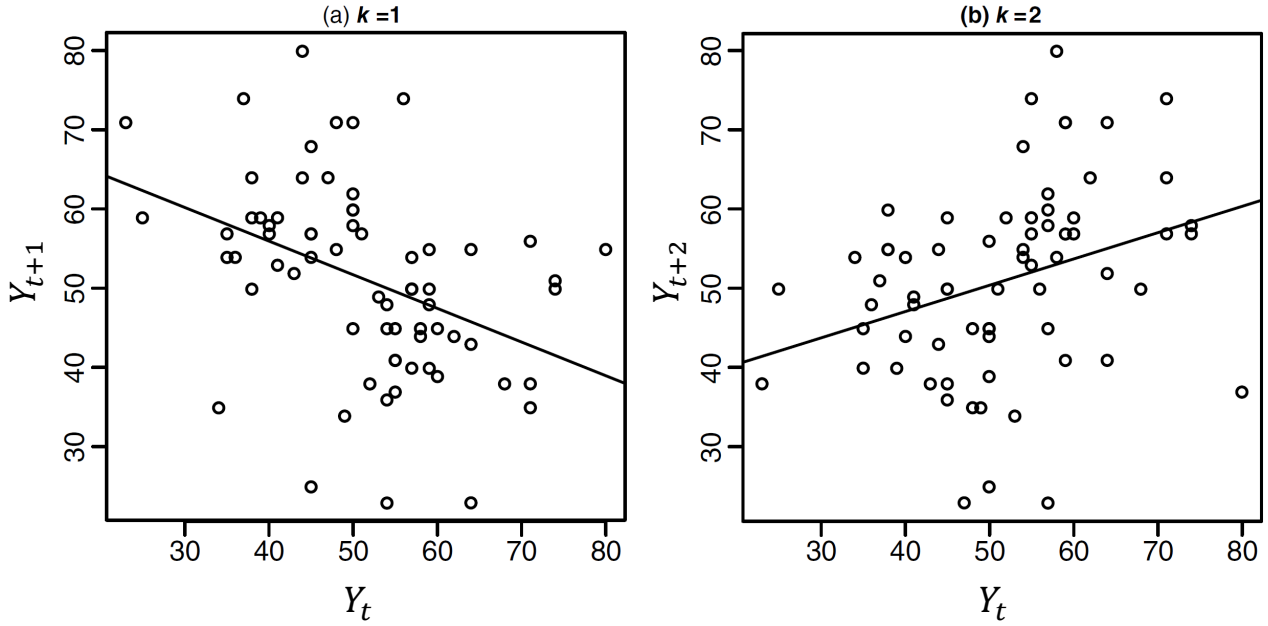
$$\gamma_{YY}(k)=\mathrm{Cov}[Y_t,Y_{t+k}]=E[(Y_t-\mu_Y)(Y_{t+k}-\mu_Y)]$$



- 自相关系数

$$\rho_{YY}(k)=\frac{E[(Y_t-\mu_Y)(Y_{t+k}-\mu_Y)]}{\sqrt{E[(Y_t-\mu_Y)^2]E[(Y_{t+k}-\mu_Y)^2]}}$$

$$=\frac{\gamma_{YY}(k)}{\gamma_{YY}(0)}=\frac{\gamma_{YY}(k)}{\sigma_Y^2}$$



弱化“平稳”性条件

- 实际中大多难以实现严平稳的要求，因此需要**放松**对不同时刻随机变量关系的约束

联合概率分布不随时间变化 → 统计量不随时间变化

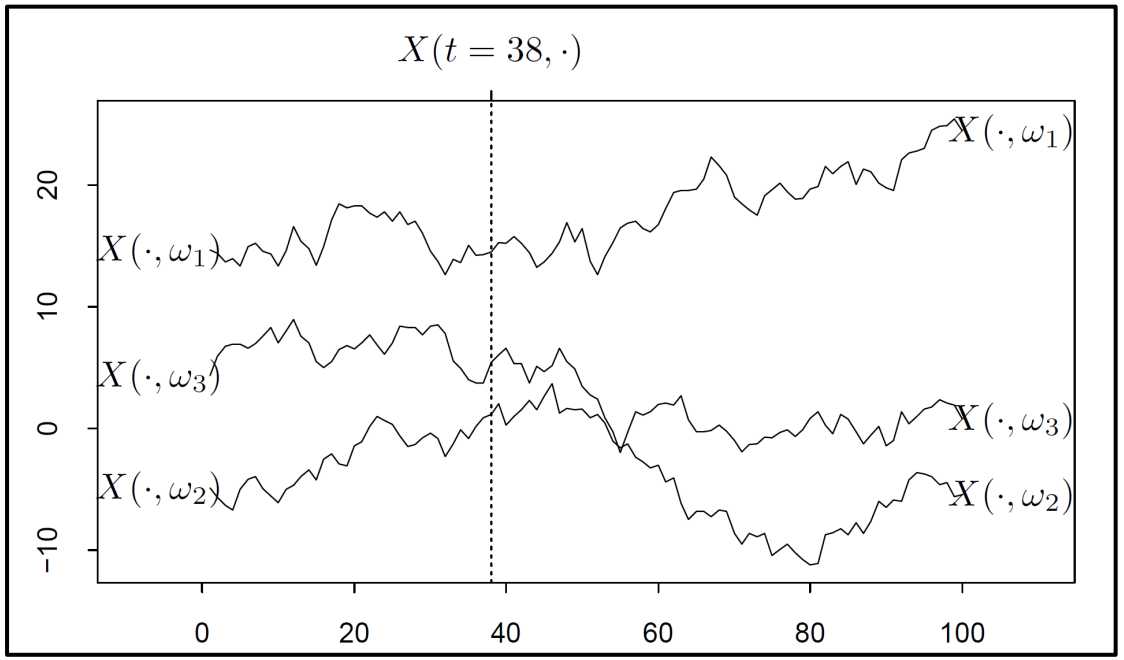
- **k阶宽平稳（weakly stationary of order k）**：随机过程的前k阶矩不随时间变化

- （二阶）宽平稳：1-2阶矩不随时间变化

$$\mu(t) = \mu, \qquad \sigma^2(t) = \sigma^2, \qquad \gamma(t_1, t_2) = \gamma(t_1 - t_2)$$

- 通常情况下，严平稳能推出宽平稳成立，反之不一定成立

随机过程和时间序列



随机过程定义为函数$Y(t,\omega)$

- 时间：$t \in T$，T为索引集（Index Set）
- 实现：$\omega \in \Omega$，Ω为 样本空间（Ensemble）

固定时间索引

$Y(t=t_0,\cdot)$

是一个随机变量

一个实现（realization）

$Y(\cdot,\omega)$

是一个时间序列

$f_{Y(t_1),\ldots,Y(t_n)}(y_1,\ldots,y_n)$为$n$维概率分布，针对$n=1,2,3\ldots$及

所有时间长度的概率分布族刻画了随机过程的性质

特殊的随机过程

- 高斯过程（Normal Process）：所有有限维度的分布均为多元高斯分布

- 马尔科夫过程（Markov Process）：条件分布只依赖于随机过程最近的状态

$$\mathrm{P}\{Y(t_n) \le y \mid Y(t_{n-1}), \cdots, Y(t_1)\}$$
$$= \mathrm{P}\{Y(t_n) \le y \mid Y(t_{n-1})\}$$

- 确定性过程（Deterministic processes）
  - 可基于历史观测进行确定性地预测（without uncertainty）
  - 如$X_t = A_1 \cos(\omega t) + A_2 \sin(\omega t)$

- 纯随机过程（Pure stochastic processes）
  - 可写为（无穷）多个无关随机变量的线性组合

$$Y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \cdots$$

$$E[\varepsilon_t] = 0, \mathrm{V}[\varepsilon_t] = \sigma^2, \sum_{i=1}^{\infty} \psi_i^2 < \infty$$

纯随机序列（白噪声）

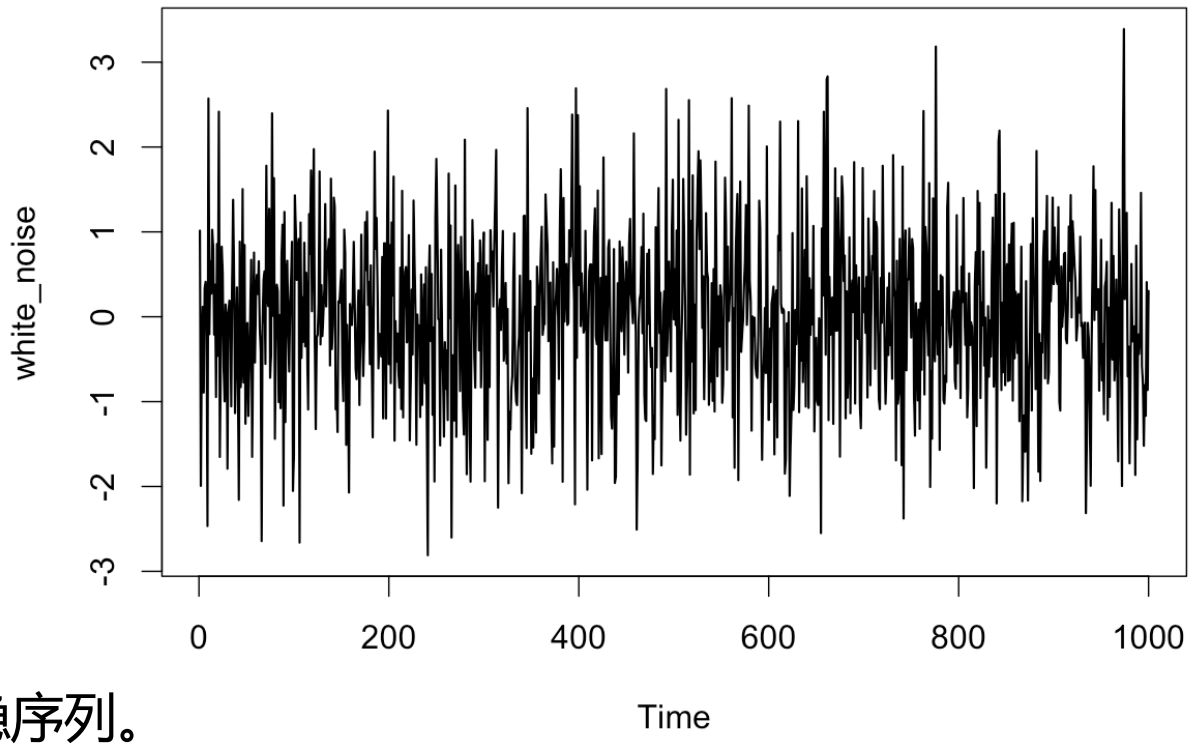
- 纯随机序列，也称为白噪声 (White Noise) 序列

$$Y_t \sim WN(0,\ \sigma^2)$$

$$E[Y_t] = 0, V[Y_t] = \sigma^2$$

$$\gamma(k) = \begin{cases} \sigma^2, & k = 0 \\ 0, & k \neq 0 \end{cases}$$

$$\rho(k) = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases}$$



- 白噪声序列一定是平稳序列, 而且是最简单的平稳序列。

若给定时间序列是白噪声序列，则无需预测

概要

1. 平稳序列的定义：

使基于有限数据的建模可行

2. 时间序列的统计量：

通过哪些统计量刻画时间序列？

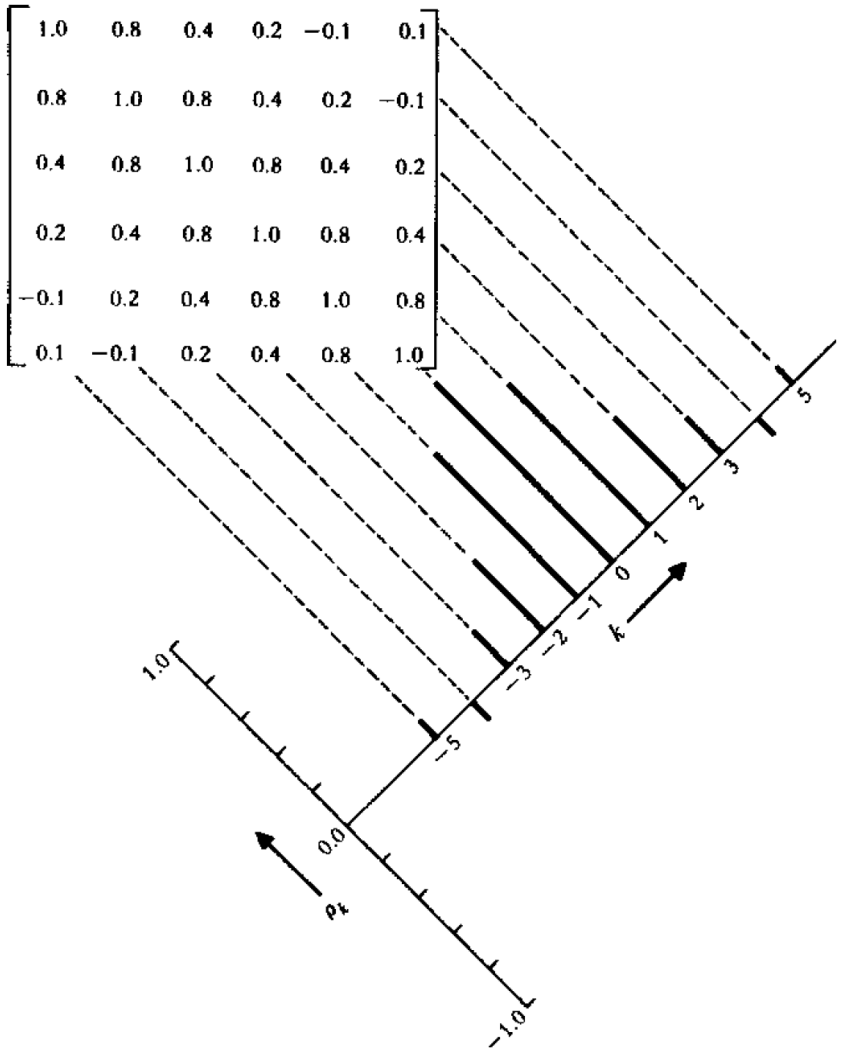
3. 时间序列分析化简：

时间序列初步建模思路

自协方差函数以及自相关函数

- 自协方差函数（Autocovariance function）
  - 自协方差系数$\gamma_{YY}(k) = \mathrm{Cov}[Y_t, Y_{t+k}]$和$k$有关
  - $\gamma_{YY}(k)$随k的变化即为自协方差函数，
  - 类似地， $\rho_{YY}(k)$为自相关函数

    (Autocorrelation function, **ACF**)



AR(1)过程

- $\{\varepsilon_t\}$为白噪声，$V[\varepsilon_t]=\sigma^2$，$AR(1)$过程定义为$Y_t=\phi Y_{t-1}+\varepsilon_t$，展开为

$$Y_t=\varepsilon_t+\phi\varepsilon_{t-1}+\phi^2\varepsilon_{t-2}+\cdots+\phi^i\varepsilon_{t-i}+\cdots$$

- 统计性质（均值、方差）

$$\mu_Y=E[Y_t]=0,\ \sigma_Y^2=V[Y_t]=\left(1+\phi^2+\cdots+\phi^{2i}+\cdots\right)\sigma^2=\frac{\sigma^2}{1-\phi^2}$$

若$|\phi|<1$则**方差收敛**，否则方差无界

- 统计性质（自协方差）
  - 若$t_1>t_2$

$$\begin{aligned}\gamma(t_1,t_2)=\mathrm{Cov}\left[Y_{t_1},Y_{t_2}\right]&=\mathrm{Cov}\left[\varepsilon_{t_1}+\phi\varepsilon_{t_1-1}+\cdots+\phi^{t_1-t_2}\varepsilon_{t_2}+\cdots,\quad \varepsilon_{t_2}+\phi\varepsilon_{t_2-1}+\cdots\right]\\&=\phi^{t_1-t_2}(1+\phi^2+\phi^4+\cdots)\sigma^2\\&=\phi^{t_1-t_2}\sigma_Y^2\end{aligned}$$

  - 若$t_1<t_2$，$\gamma(t_1,t_2)=\phi^{t_2-t_1}\sigma_Y^2$
  - 综上，$\gamma(t_1,t_2)=\phi^{|t_1-t_2|}\sigma_Y^2$

$|\phi|<1$时时间序列$\{Y_t\}$为宽平稳

宽平稳条件下的随机过程统计量

一般情况下

- **自协方差**函数（Autocovariance function）

$$\gamma_{YY}(t_1,t_2)=\gamma(t_1,t_2)$$
$$=\mathrm{Cov}[Y(t_1),Y(t_2)]$$

- **自相关**函数（Autocorrelation function, ACF）

$$\rho_{YY}(t_1,t_2)=\rho(t_1,t_2)$$
$$=\frac{\gamma_{YY}(t_1,t_2)}{\sqrt{\sigma^2(t_1)\sigma^2(t_2)}}$$

宽平稳情况下

- **自协方差**函数（Autocovariance function）

$$\gamma_{YY}(k)=\mathrm{Cov}[Y(t),Y(t+k)]$$

只和**时间间隔（延迟）**k有关

- **自相关**函数（Autocorrelation function, ACF）

$$\rho_{YY}(k)=\frac{\gamma_{YY}(k)}{\gamma_{YY}(0)}=\frac{\gamma_{YY}(k)}{\sigma_Y^2}$$

$$\rho_{YY}(0)=1$$

平稳条件下自协方差函数的性质

- 平稳随机过程$\{Y(t)\}$自协方差函数$\gamma(k)$的性质
  - 【对称性】$\gamma(k)=\gamma(-k)$

$$\gamma(k)=\mathrm{Cov}[Y(t),Y(t+k)]=\mathrm{Cov}[Y(t+k),Y(t)]=\mathrm{Cov}[Y(t),Y(t-k)]=\gamma(-k)$$

  - 【规范性】$|\gamma(k)|\le\gamma(0)$

由于$\mathrm{V}[\lambda_1Y(t)+\lambda_2Y(t+k)]\ge 0$，因此

$$(\lambda_1^2+\lambda_2^2)\gamma(0)+2\lambda_1\lambda_2\gamma(k)\ge 0$$

若$\lambda_1=\lambda_2=1$，有$\gamma(k)\ge-\gamma(0)$，若$\lambda_1=1,\lambda_2=-1$，则$\gamma(k)\le\gamma(0)$

  - 【非负定性】$\forall y,n,t$，二次型$\sum_{i=1}^{n}\sum_{j=1}^{n}y_iy_j\gamma\left(t_i-t_j\right)\ge 0$
  - 【连续性】若$\gamma(k)$在$k=0$处连续则$\gamma(k)$对任意k连续

平稳条件下自相关函数的性质

- 平稳随机过程$\{Y(t)\}$自相关函数$\rho(k)$的性质
  - 【对称性】$\rho(k) = \rho(-k)$
  - 【规范性】$|\rho(k)| \leq 1$
  - 【非负定性】$\forall n$，矩阵为非负定阵

$$\begin{bmatrix} \rho(0) & \rho(1) & \cdots & \rho(n-1) \\ \rho(1) & \rho(0) & \cdots & \rho(n-2) \\ \vdots & \vdots & \cdots & \vdots \\ \rho(n-1) & \rho(n-2) & \cdots & \rho(0) \end{bmatrix}$$

  - 【非唯一性】一个平稳时间序列一定唯一决定了它的自相关函数，但一个自相关函数未必唯一对应着一个平稳时间序列

宽平稳条件下的随机过程统计量

一般情况下

- 考虑随机过程$\{X(t)\}$和$\{Y(t)\}$
- **协方差**函数（cross-covariance function）

$$\gamma_{YX}(t_1,t_2)=\mathrm{Cov}[Y(t_1),X(t_2)]$$
$$=\mathrm{E}\big[\big(Y(t_1)-\mu_Y(t_1)\big)\big(X(t_2)-\mu_X(t_2)\big)\big]$$

- **相关**函数（cross-correlation function, CCF）

$$\rho_{YX}(t_1,t_2)=\frac{\gamma_{YX}(t_1,t_2)}{\sqrt{\sigma_Y^2(t_1)\sigma_X^2(t_2)}}$$

宽平稳情况下

- 考虑随机过程$\{Y(t)\}$和$\{X(t)\}$
- **协方差**函数（cross-covariance function）

$$\gamma_{YX}(k)=\mathrm{Cov}[Y(t),X(t+k)]$$

- **相关**函数（cross-correlation function, CCF）

$$\rho_{YX}(k)=\frac{\gamma_{YX}(k)}{\sqrt{\gamma_{YY}(0)\gamma_{XX}(0)}}=\frac{\gamma_{YX}(k)}{\sigma_Y\sigma_X}$$

AR(1)过程的自相关函数

- $\{\varepsilon_t\}$为白噪声，$V[\varepsilon_t]=\sigma^2$，$AR(1)$过程为

$$Y_t=\phi Y_{t-1}+\varepsilon_t=\varepsilon_t+\phi\varepsilon_{t-1}+\phi^2\varepsilon_{t-2}+\cdots+\phi^i\varepsilon_{t-i}+\cdots$$

- 统计性质（自协方差）

$$\gamma(k)=\gamma(-k)=\mathrm{Cov}[Y_t,Y_{t-k}]=\mathrm{Cov}[\phi Y_{t-1}+\varepsilon_t,Y_{t-k}]$$
$$=\phi\,\mathrm{Cov}[Y_{t-1},Y_{t-k}]=\phi\gamma(k-1)=\phi^2\gamma(k-2)=\cdots$$

因此$\gamma(k)=\phi^{|k|}\gamma(0)$, 自相关函数为

$$\rho(k)=\phi^{|k|},|\phi|<1$$

ϕ表示时间序列的记忆性

平稳条件下统计量的估计

- 常数均值

$$\mu(t) = E[Y(t)] = \int_{-\infty}^{\infty} y f_{Y(t)}(y) dy$$

$$= \mu$$

- 均值估计

$$\hat{\mu} = \frac{1}{n}\sum_{t=1}^{n} y_t$$

- 无偏性

$$E[\hat{\mu}] = \mu$$

- 计算$\hat{\mu}$的精确程度（方差）

$$\mathrm{var}[\hat{\mu}] = \frac{1}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\gamma(t-s)$$

$$= \boxed{\frac{\gamma(0)}{n}}\left[1 + 2\sum_{k=1}^{n-1}\left(1-\frac{k}{n}\right)\rho(k)\right]$$

独立同分布序列
均值估计的方差

$$\approx \frac{\gamma(0)}{n}\left[1 + 2\sum_{k=1}^{n-1}\rho(k)\right] \ (n\text{较大})$$

平稳条件下统计量的估计

- 时间序列的统计量（真实值）
  - 均值（序列）

$$\mu(t) = E[Y(t)] = \int_{-\infty}^{\infty} y f_{Y(t)}(y) dy$$

  - 自协方差

$$\gamma_{YY}(t_1, t_2) = \gamma(t_1, t_2) = \mathrm{Cov}[Y(t_1), Y(t_2)]$$
$$= E\left[\left(Y(t_1) - \mu(t_1)\right)\left(Y(t_2) - \mu(t_2)\right)\right]$$

- 时间序列的统计量（估计值）
  - 均值（序列）

$$\hat{\mu} = \frac{1}{n}\sum_{t=1}^{n} y_t$$

  - 延迟k自协方差函数

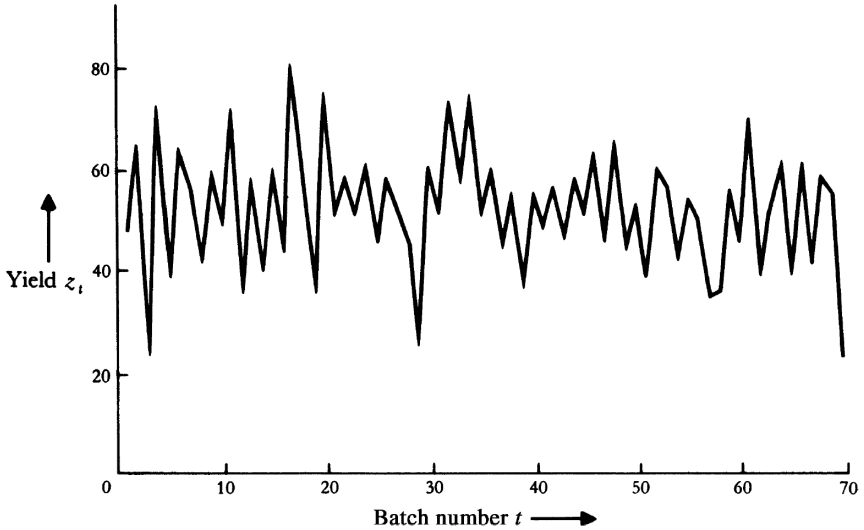
$$\hat{\gamma}(k) = \frac{1}{n}\sum_{t=1}^{n-k}(y_t - \hat{\mu})(y_{t+k} - \hat{\mu}), 0 < k < n$$

  - 方差 $(k = 0)$

$$\gamma(0) = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{\mu})$$

平稳性的意义：极大减少随机变量的个数，增加待估变量的样本量

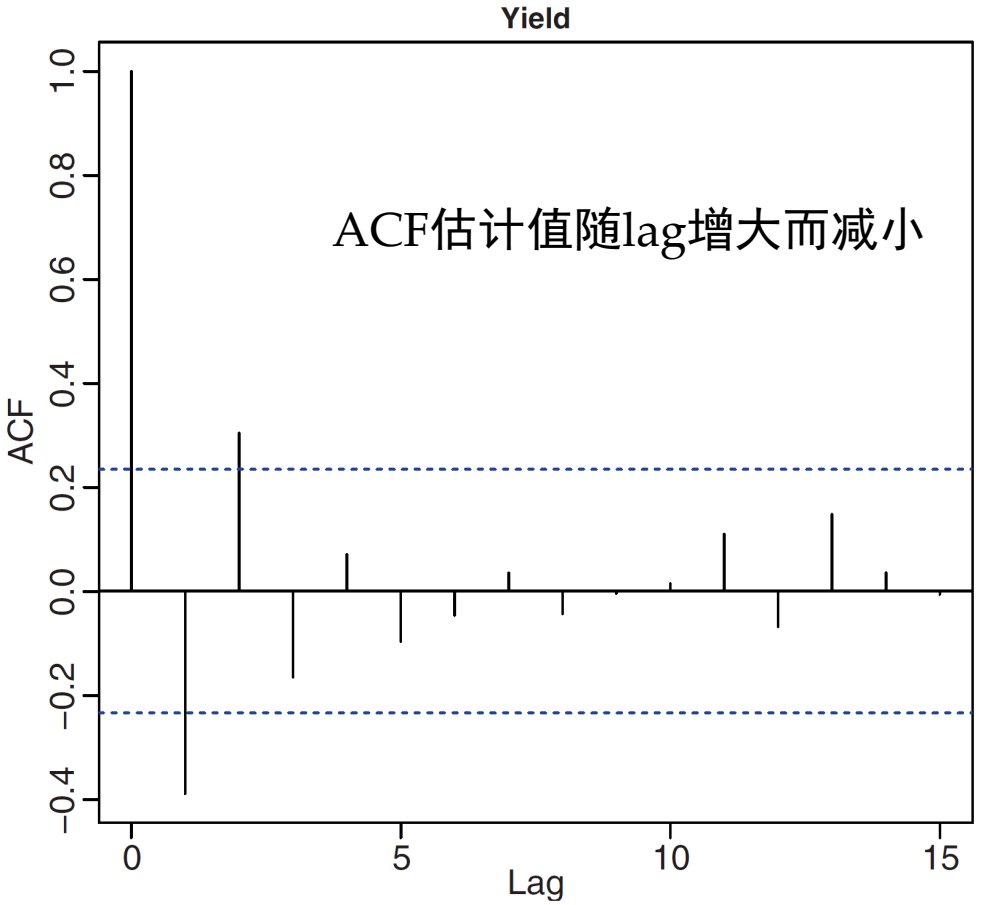
自协方差函数估计



原始数据

k	$\hat{\rho}_k$	k	$\hat{\rho}_k$	k	$\hat{\rho}_k$
1	−0.39	6	−0.05	11	0.11
2	0.30	7	0.04	12	−0.07
3	−0.17	8	−0.04	13	0.15
4	0.07	9	0.00	14	0.04
5	−0.10	10	0.01	15	−0.01

ACF估计值



ACF图

自协方差函数估计的标准差

- （当lag较大时）对所估计的ACF进行（是否是0的）检测。Bartlett [1946]给出对于平稳的高斯过程估计自相关系数的方差

$$V[\hat{\rho}_k] \simeq \frac{1}{N}\sum_{v=-\infty}^{\infty}\left(\rho_v^2 + \rho_{v+k}\rho_{v-k} - 4\rho_k\rho_v\rho_{v-k} + 2\rho_v^2\rho_k^2\right)$$

- 若$\rho_k = \phi^{|k|}, -1 < \phi < 1$，自相关函数ACF指数衰减

$$V[\hat{\rho}_k] \simeq \frac{1}{N}\left[\frac{(1+\phi^2)(1-\phi^{2k})}{1-\phi^2} - 2k\phi^{2k}\right]$$

且

$$\mathrm{var}[\hat{\rho}_1] \simeq \frac{1}{N}(1-\phi^2)$$

- 若当$v > q$有$\rho_v = 0$，则

$$V[\hat{\rho}_k] \simeq \frac{1}{N}\left(1 + 2\sum_{v=1}^{q}\rho_v^2\right), \qquad k > q$$

large-lag standard error

自协方差函数估计的标准差

- 针对$n = 200$的序列，已知$\rho_1 = -0.4, \rho_k = 0\ (k \geq 2)$，计算出自相关系数估计值
- 假设序列为白噪声序列，$q = 0$，则

$$\mathrm{var}[\hat{\rho}_k] \simeq \frac{1}{n} = \frac{1}{200} = 0.005, \qquad k > 0$$

标准差约0.07.

- $\rho_1 \neq 0$，当$k \geq 2$时，$\hat{\rho}_k$是否为0?
- $q = 1$代入$\hat{\rho}_1$和ρ_1，large-lag variance为

$$\mathrm{var}[\hat{\rho}_k] \simeq \frac{1}{200}[1 + 2(-0.38)^2] = 0.0064, \qquad k > 1$$

标准差为0.08, $k > 1$所估计的自相关系数相比于该标准差较小，因此认为$\rho_1 \neq 0$，$\rho_k = 0\ (k \geq 2)$

大于5倍标准差，显著非0

k	$\hat{\rho}_k$	k	$\hat{\rho}_k$
1	−0.38	6	0.00
2	−0.08	7	0.00
3	0.11	8	0.00
4	−0.08	9	0.07
5	0.02	10	−0.08

概要

1. 平稳序列的定义：
使基于有限数据的建模可行

2. 时间序列的统计量：
通过哪些统计量刻画时间序列？

3. 时间序列分析化简：
时间序列初步建模思路

Wold分解定理

Herman Wold
1938

平稳序列分解定理：任意**离散平稳**时间序列$\{Y_t\}$都可以分解为两个不相关的**平稳**序列之和，记$\{V_t\}$为**确定性**（deterministic) 平稳序列，$\{\xi_t\}$ 为**随机性** (stochastic) 平稳序列，则$Y_t = V_t + \xi_t$

- Wold证明$\{V_t\}$可以等价表达为**历史序列值的线性函数**

$$V_t = \sum_{j=1}^{\infty} \phi_j Y_{t-j}$$

- 确定性序列$\{V_t\}$是序列中当期波动可以由其历史序列值解读的部分。

- Wold证明随机序列$\{\xi_t\}$可以等价表达为

$$\xi_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

式中：$\{\varepsilon_t\}$称为新息过程（innovation process），是每个时期加入的新的随机信息，相互无关，不可预测。通常假定$\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$，$\forall t \geq 0$。且有$\psi_0 = 1$，$\sum_{j=0}^{\infty} \psi_j^2 < \infty$。

Cramer分解定理

通过差分使序列平稳

Harald Cramer
瑞典统计学家和保险精算学家

任何一个序列的波动都可以视为同时受到了确定性影响和随机性影响的综合作用：

$$Y_t = \sum_{j=1}^{d} \beta_j t^j + \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j}$$

- 一阶差分

$$\nabla Y_t = Y_t - Y_{t-1}$$

- p阶差分

$$\nabla^p Y_t = \nabla^{p-1} Y_t - \nabla^{p-1} Y_{t-1}$$

- k步差分

$$\nabla_k Y_t = Y_t - Y_{t-k}$$

- Cramer分解定理在理论上保证了**适当阶数的差分**一定可以充分提取确定性信息

$$\nabla^d \sum_{j=0}^{d} \beta_j t^j = c$$

c为某一常数。差分运算的实质是使用**自回归**的方式提取确定性信息

$$\nabla^d Y_t = (1-B)^d Y_t = \sum_{i=0}^{d} (-1)^i C_d^i Y_{t-i}$$

延迟算子

- 延迟算子类似于一个时间指针，当前序列值乘以一个延迟算子，就相当于把当前序列值的时间向过去拨了一个时刻
- 记**B**为延迟算子，有

$$Y_{t-1} = \mathbf{B}Y_t$$

$$Y_{t-2} = \mathbf{B}^2 Y_t$$

$$\vdots$$

$$Y_{t-p} = \mathbf{B}^p Y_t, \forall p \geq 1$$

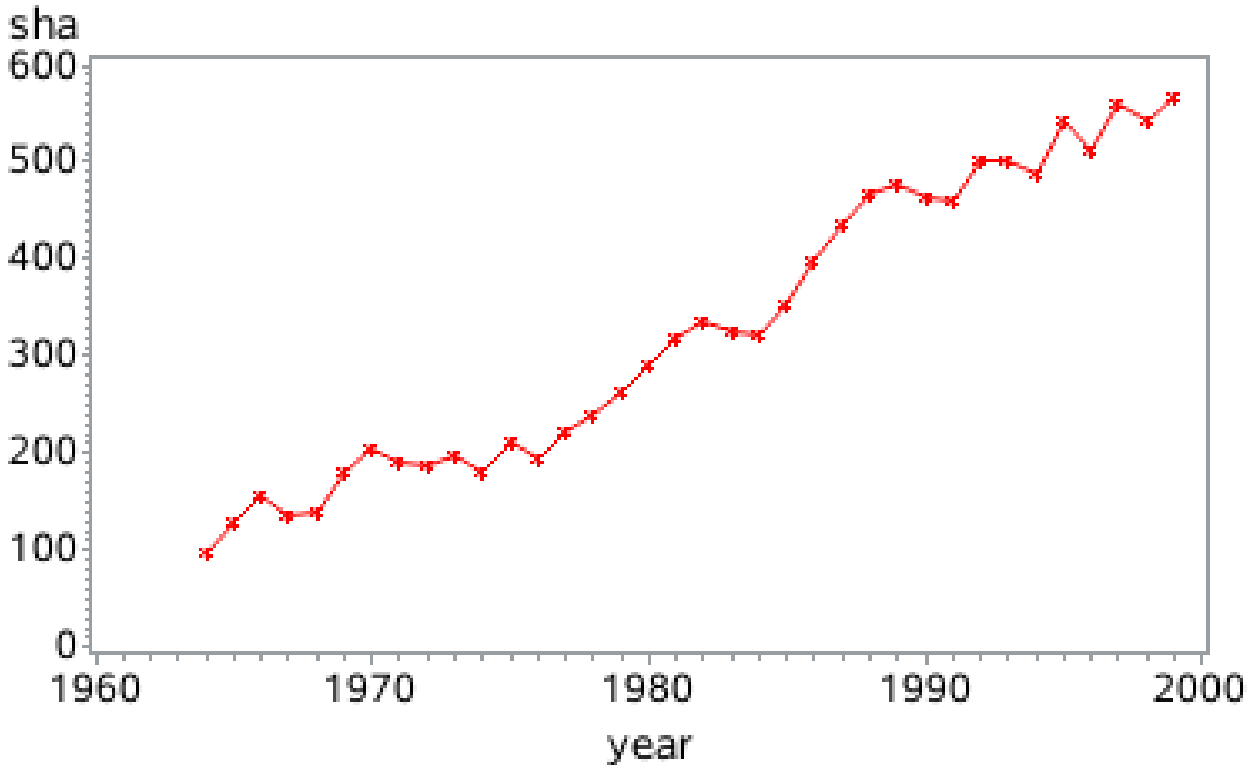
- $\mathbf{B}^{-1}$使序列向未来移动
- $\mathbf{B}^0 = 1$
- $\mathbf{B}(c \cdot Y_t) = c \cdot \mathbf{B}(Y_t) = c \cdot Y_{t-1}$, c为任意常数
- $\mathbf{B}(Y_t \pm Y_t) = Y_{t-1} \pm Y_{t-1}$
- $\mathbf{B}^n Y_t = Y_{t-n}$
- $(1-\mathbf{B})^n = \sum_{i=0}^{n}(-1)^i C_n^i \mathbf{B}^i$， 其中

$$C_n^i = \frac{n!}{i!(n-i)!}$$

差分方式的选择

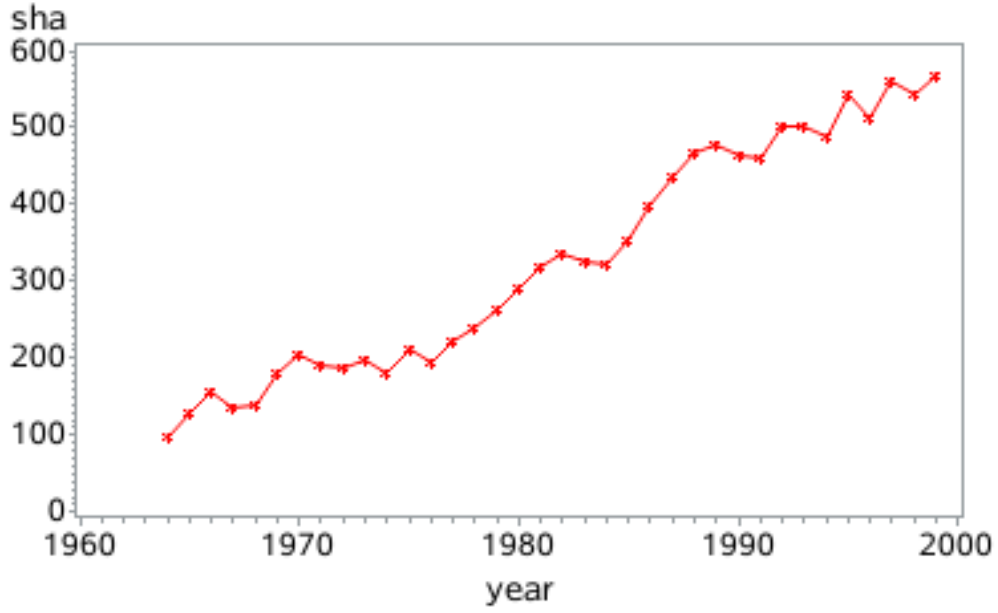
- 序列蕴含着显著的**线性**趋势，通常一阶差分就可以实现趋势平稳
- 序列蕴含着**曲线**趋势，通常低阶（二阶或三阶）差分就可以提取出曲线趋势的影响
- 对于蕴含着**固定周期**的序列，进行步长为周期长度的差分运算，通常可以较好地提取周期信息

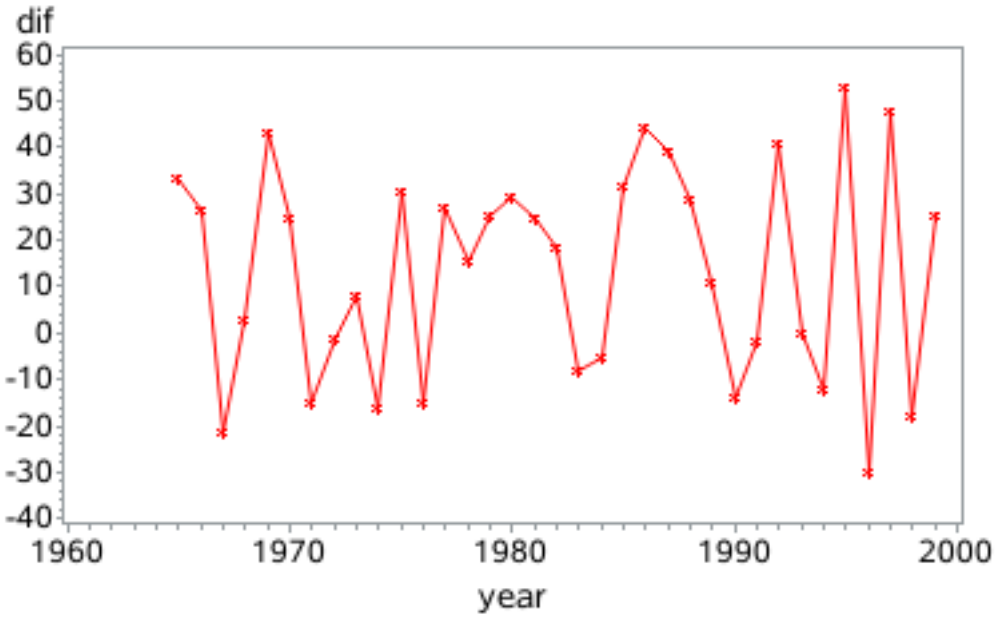
提取1964-1999年中国纱年产量序列中的确定性信息



差分提取时间序列中确定性信息

1964-1999年中国纱年产量序列



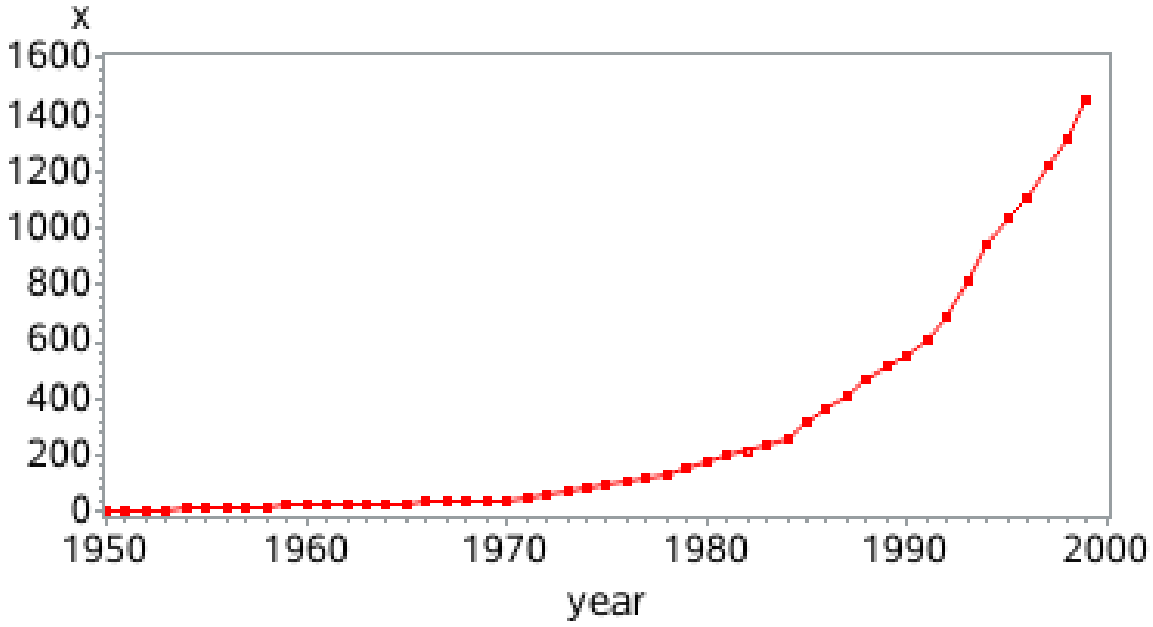


中国纱产量序列蕴涵着显著的线性递增趋势。对该序列进行**1阶差分**提取线性趋势信息

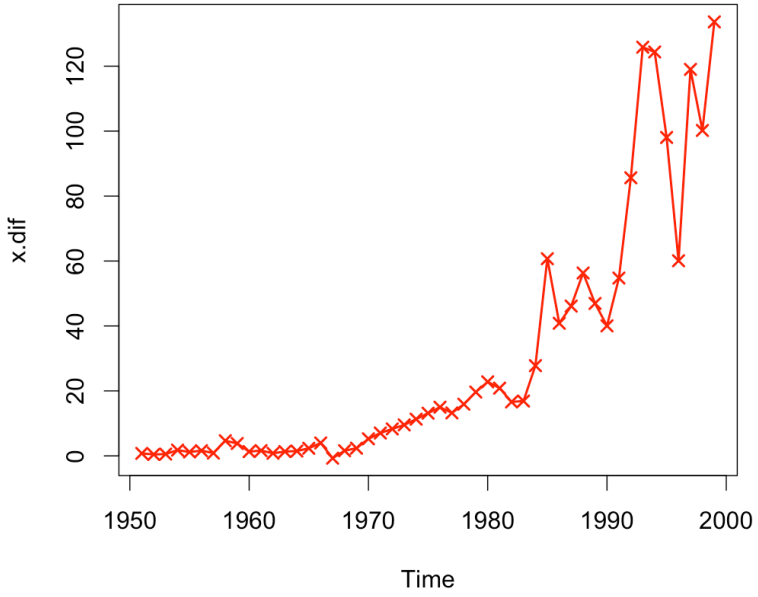
1阶差分后序列呈现出非常平稳的波动特征，这说明1阶差分运算非常成功地从原序列中提取出线性趋势。

差分提取时间序列中确定性信息

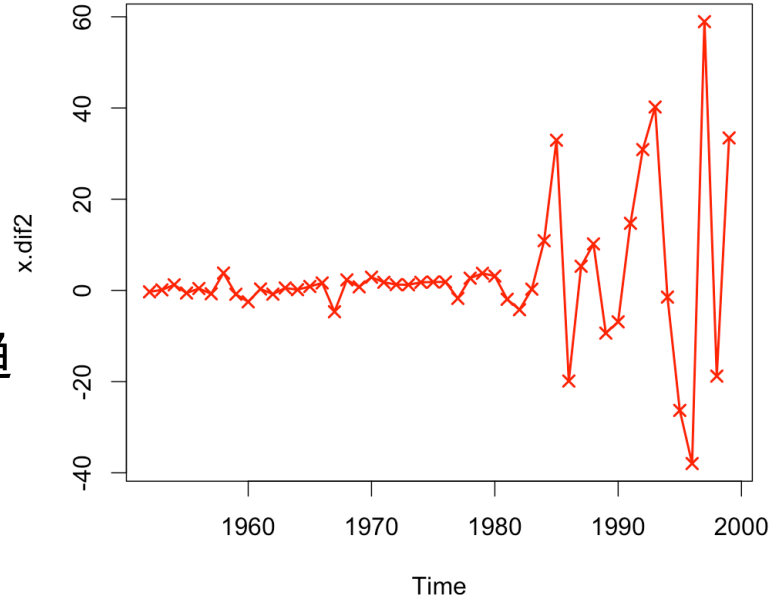
1950—1999年北京市民用车辆拥有量序列



1阶差分没有实现趋势平稳

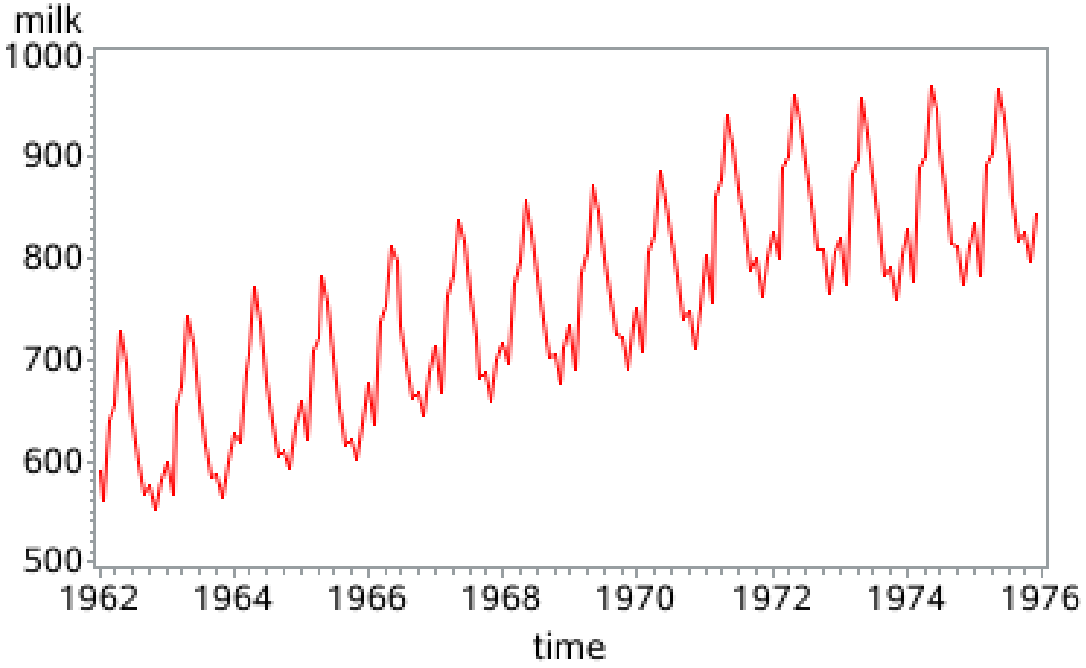


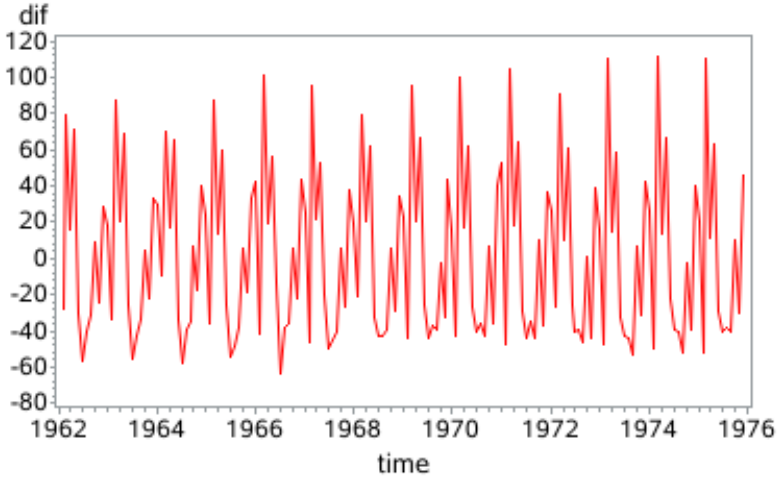
2阶差分实现趋势平稳，但随机波动不平稳



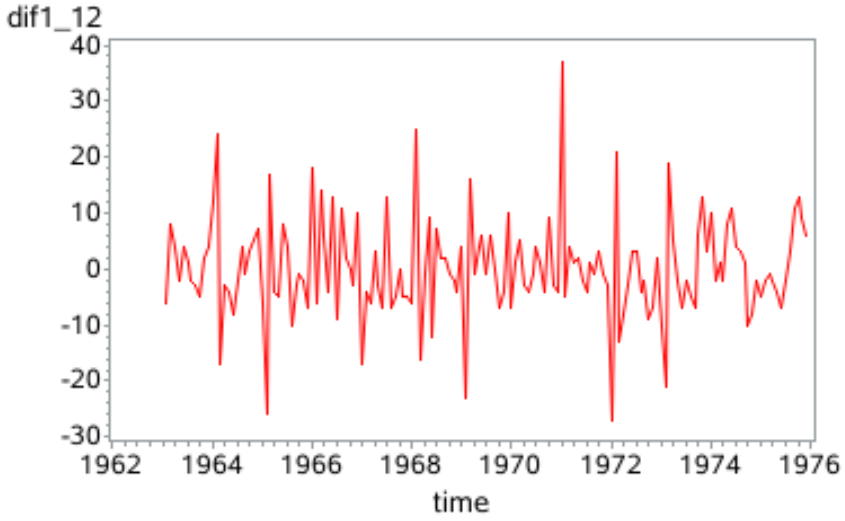
差分提取时间序列中确定性信息

1962年1月至1975年12月平均每头奶牛的月产奶量序列





1 阶差分后线性递增信息被提取，残留稳定的**季节波动**和**随机波动**



1阶12步差分：周期差分可以提取**周期**信息

过差分

- 虽然足够多次的差分运算可以充分地提取原序列中的非平稳确定性信息，但每次差分都会有信息的损失，应当避免过度差分的现象 。

- 时间序列

$$Y_t = \beta_0 + \beta_1 t + a_t$$

$$E(a_t) = 0, Var(a_t) = \sigma^2, Cov(a_t, a_{t-i}) = 0, \forall i \geq 1$$

- 一阶差分
- 平稳

$$\nabla Y_t = Y_t - Y_{t-1} = \beta_1 + a_t - a_{t-1}$$

- 方差小

$$Var(\nabla Y_t) = Var(a_t - a_{t-1}) = 2\sigma^2$$

- 二阶差分（过差分）

平稳

$$\nabla^2 Y_t = \nabla Y_t - \nabla Y_{t-1} = a_t - 2a_{t-1} + a_{t-2}$$

方差大

$$Var(\nabla^2 Y_t) = Var(a_t - 2a_{t-1} + a_{t-2}) = 6\sigma^2$$