

第五章 气候变化趋势分析

研究对象 气候时间序列 x_t : 随时间变化的一列气候数据。

如年降水总量、月海表温度、日最高(低)气温、积雪深度及日数、台风强度及路径等序列。

基本特点

- ① 数据取值随时间变化
- ② 每一时刻取值的随机性
- ③ 前后时刻数据之间存在相关性、持续性
- ④ 序列整体有上升或下降趋势，或呈周期振荡
- ⑤ 某一时刻数据取值出现转折或突变
- 时间序列不是同时满足这些特性的

公式

任一气候时间序列 x_t 可看成由以下几个分量构成:

$$x_t = H_t + P_t + C_t + S_t + a_t \quad H_t \text{ 气候趋势分量} \\ C_t \text{ 循环变化分量}$$

P_t 气候序列固有的周期性变化

S_t 平稳时间序列分量 a_t 随机扰动项

分解时间序列: 要分离 H_t 的常用做法, 采用年、月、季节总量或平均值等来构造气候时间序列。这样可以消除 P_t , 再通过统计处理消除或削弱 C_t 和 a_t 。

5.1 线性倾向估计

5.1.1 方法介绍

方法介绍 用 x_i 表示样本量为 n 的某一气候变量, 用 t_i 表示 x_i 所对应的时间, 建立 x_i 与 t_i 之间的一元线性回归关系:

$$\hat{x}_i = a + bt_i \quad (i = 1, 2, \dots, n)$$

含义 用一条合理的直线 \hat{x}_i 表示 x 与其时间 t 之间的关系。 a, b 是回归系数, a, b 可用最小二乘法进行估计。

计算方法 $a = \bar{x} - b\bar{t}$ $b = \frac{\sum_{i=1}^n x_i t_i - n\bar{x}\bar{t}}{\sum_{i=1}^n t_i^2 - n\bar{t}^2} = \frac{s_{xt}}{s_t^2}$ **相关系数:** $r = \frac{s_t}{s_x} \times b = \sqrt{\frac{\sum_{i=1}^n t_i^2 - n\bar{t}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}} \times b$

5.1.2 计算步骤

第一步 对变量 x_i 构造其对应的时间序列 t_i 。可以是年份, 也可以是自然数序列等。

第二步 根据公式算出回归系数 b , 回归常数 a 及相关系数 r 。

第三步 将 a, b 的值代入回归方程式, 求出回归计算值。

5.1.3 计算结果分析

回归系数 b 又称为**倾向值**, b 的符号说明了气候变量 x 的**趋势倾向**。当 $b > 0$, x 是呈上升趋势; 当 $b < 0$, 则为下降。 b 值的大小反映了上升或下降的速率, b 的绝对值越大, 表明直线越倾斜。

相关系数 r 相关系数反映了 x 与 t 间的**密切程度**。当 $r = 0$, 说明 x 的变化与时间无关。 $r > 0$ 说明 x 随时间 t 的增加, x 呈上升趋势; $r < 0$, 则相反。 r 的绝对值越大, 说明 x 与时间的关系越密切, 这与 b 所反映的意义是一致的。

显著性检验 要判断变化趋势的程度是否显著, 可以对 b, r 进行显著性检验 (**F检验与t检验**)。

实例

使用线性倾向估计分析华北地区 1951-1995 年夏季干旱指数的变化趋势: 分析代表华北整个区域干旱状况的干旱指数的变化趋势。这里 $n = 45$, x_i 为干旱指数。计算出 $a = 40.76$, $b = -0.0182$, 相关系数 $r = -0.3395$ 。计算结果表明, 从总体上考察, 该地区夏季干旱指数呈下降趋势, 相关系数 $|r| > r_{0.05} = 0.2875$, 表明这种下降趋势在 $\alpha = 0.05$ 显著性水平上是显著的。

5.2 滑动平均

5.2.1 方法概述

方法概述 滑动平均是趋势拟合技术最基础的方法，它相当于**低通滤波器**，滤除了高频成分。用确定时间序列的平滑值来显示变化趋势。对样本量为n的气候序列x，其滑动平均序列 \hat{x}_j 表示为：

$$\hat{x}_j = \frac{1}{k} \sum_{i=1}^k x_{i+j-1} \quad (j = 1, 2, \dots, n - k + 1)$$

式中k为滑动长度，一般取奇数（如果进行三点滑动平均，最终得到n-2样本量的序列）。

案例

如：对一个n = 23的时间序列进行k = 3点平滑，则新序列为： $\hat{x}_j = \frac{1}{3} \sum_{i=1}^3 x_{i+j-1}$ ($j = 1, 2, 3, \dots, 21$)，即： $\hat{x}_1 = \frac{1}{3}(x_1 + x_2 + x_3), \hat{x}_2 = \frac{1}{3}(x_2 + x_3 + x_4), \dots, \hat{x}_{21} = \frac{1}{3}(x_{21} + x_{22} + x_{23})$ 等权重

如果考虑年代际的变化，k可以取11年。除此以外，还可以进行频谱分析。

含义 经过滑动平均后，序列中**短于滑动长度的周期大大削弱**，体现出变化趋势来。但经过这种滑动平均后的新序列比原气候时间序列短，原序列两头的信息不能体现。用这种方法求得的是各个时刻的**趋势值**，而不是具体的数字表达式。

检验 滑动平均后也是一个时间序列，可以正常使用自由度更低的显著性检验方法。

5.2.2 计算步骤

人工计算 n个数据可以得到n-k+1的平滑值。

编程计算 先用前k个数据求和，得到一个数字；然后依次用这个数字减去平均时段的第一个数字，并加上第k+1个数据，再用求出的值除以k；循环这样的过程计算出第2个到第n-k+1个平滑值。第一个平滑值就是前k个数据的平均值。

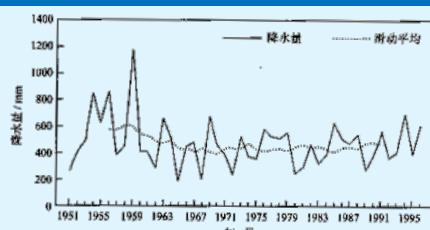
5.2.3 计算结果分析

分析方法 主要从滑动平均序列曲线图来诊断其变化趋势。例如：看其演变趋势有几次明显的波动，是呈上升还是下降趋势。

实例 1

计算北京1951-1996年夏季降水量的11年滑动平均。样本量n = 46，滑动平均后得到36个平滑值。图中较光滑的曲线即为滑动平均曲线。

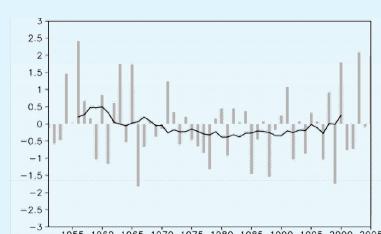
可以看出，20世纪50年代中期至60年代末，北京夏季降水量呈逐渐下降走势。20世纪70年代初降至低点后变化平缓，处于少雨阶段，并持续到今，虽有小的波动。但没有出现明显的上升或下降趋势。



实例 2

右图是环河流域夏季降水标准化距平及其11年平滑图。可以看到明显的年代际变迁，且整体呈现下降趋势。

由于我们损失了k-1个数据，因此滑动平均图两端必然比原有时间序列短。但是，在论文中常常看到补全的行为（用原始值直接衔接），这是错误的做法。这么做可能导致错误的结论推导。



5.3 累积距平

5.3.1 方法概述

方法概述 累积距平是由曲线直观判断变化趋势的方法。对于序列 x_i , 其某一时刻 t 的累积距平表示为:

$$\hat{x}_t = \sum_{i=1}^t (x_i - \bar{x}) \quad t = 1, 2, \dots, n$$

第一个值为距平值, 第二个值为前两个距平值之和, 最后一个值为所有距平值之和, 必然等于零。

5.3.2 计算步骤

第一步 计算出 x 的均值

第二步 根据上式逐一计算出各个时刻的累积距平值

注意 **最后一个累积距平值为 0** (全部距平的和)。

x_i	4	3	1	5	7	4
x_{di}	0	-1	-3	1	3	
x_{li}	0	-1	-4	-3	0	

计算实例

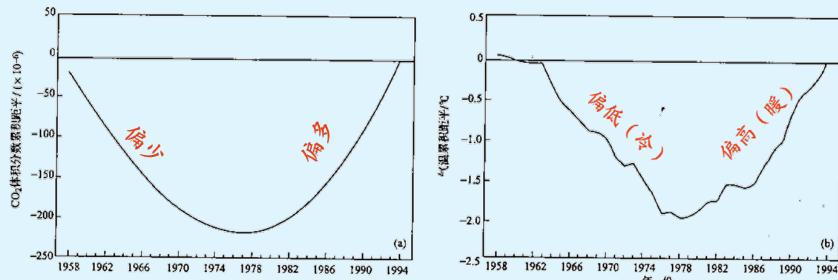
5.3.3 计算结果分析

分析要点

- ① 累积距平曲线**呈上升趋势**, 表示**累积距平值增大(正距平)**, 气候变量**以偏多(高)状态为主**。
- ② **呈下降趋势**, 表示**累积距平值减小(负距平)**, 气候变量**以偏少(低)状态为主**。
- ③ 从**曲线明显的上下起伏**, 可以判断其长期显著的**演变趋势及持续性变化**, 甚至还可以判断出发生突变的大致时间。从**曲线小的波动**可以考察其短期的距平值变化。

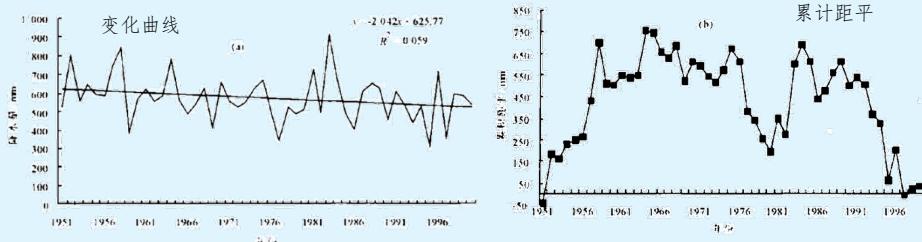
实例 1

1958-1994 年全球二氧化碳浓度 (a) 和全球气温 (b) 的累积距平曲线。



实例 2

西安地区 1951-2000 年降水量的累积距平曲线 (左图) 表明: 1951-1958 年、1980-1984 年曲线趋势上升, 亦为降水量偏多。1959-1974 年曲线趋势变化小, 降水量基本正常。从 1975-1980 年、1991-2000 年曲线趋势下降, 亦为降水量偏少。



5.4 五、七和九点二次平滑

方法引入

滑动平均是等权重平均，而实际情况下我们往往希望越近的样本权重越高，由此引入多点二次平滑。对于任何资料，其都是多种不同时间尺度因子耦合作用的结果，我们可以简单分为年代信号和年际信号，我们先去除年代信号，然后对去除后的序列作二次平滑，便可以得到年代信号。

天气尺度波动一般在3~8天，可以使用。滤波分为低通、高通和带通三种，带通滤波可以将特定

5.4.1 方法概述

方法概述 对时间序列作五点二次、七点二次和九点二次平滑，起到**低通滤波**作用，以展示出变化趋势。

优点 可以克服滑动平均**削弱过多波幅**的不足（存在权重的比例）。

公式推导

对时间序列 x ，用二次多项式拟合： $\hat{x}_j = a_0 + a_1 t_j + a_2 t_j^2$ 以五点为例，其误差为： $y = \sum_{j=i-2}^{i+2} (\hat{x}_j - x_j)^2 = \sum_{j=i-2}^{i+2} (a_0 + a_1 t_j + a_2 t_j^2 - x_j)^2$ 要使拟合效果最好，即是使 y 最小。根据最小二乘法原理确定系数 a_0 、 a_1 、 a_2 ，即：

$$\begin{cases} \frac{\partial y}{\partial a_0} = 0 \\ \frac{\partial y}{\partial a_1} = 0 \\ \frac{\partial y}{\partial a_2} = 0 \end{cases} \Rightarrow \begin{cases} \frac{\partial y}{\partial a_0} = 2 \sum_{j=i-2}^{i+2} (\alpha_0 + \alpha_1 t_j + \alpha_2 t_j^2 - x_j) = 0 \Rightarrow 5\alpha_0 + \alpha_1 \sum_{j=i-2}^{i+2} t_j + \alpha_2 \sum_{j=i-2}^{i+2} t_j^2 - \sum_{j=i-2}^{i+2} x_j = 0(1) \\ \frac{\partial y}{\partial a_1} = 2 \sum_{j=i-2}^{i+2} (\alpha_0 + \alpha_1 t_j + \alpha_2 t_j^2 - x_j) t_j = 0 \Rightarrow \alpha_0 \sum_{j=i-2}^{i+2} t_j + \alpha_1 \sum_{j=i-2}^{i+2} t_j^2 + \alpha_2 \sum_{j=i-2}^{i+2} t_j^3 - \sum_{j=i-2}^{i+2} x_j t_j = 0(2) \\ \frac{\partial y}{\partial a_2} = 0 \quad \frac{\partial y}{\partial a_2} = 2 \sum_{j=i-2}^{i+2} (\alpha_0 + \alpha_1 t_j + \alpha_2 t_j^2 - x_j) t_j^2 = 0 \Rightarrow \alpha_0 \sum_{j=i-2}^{i+2} t_j^2 + \alpha_1 \sum_{j=i-2}^{i+2} t_j^3 + \alpha_2 \sum_{j=i-2}^{i+2} t_j^4 - \sum_{j=i-2}^{i+2} x_j t_j^2 = 0(3) \end{cases}$$

令 $t_{i-2} = -2, t_{i-1} = -1, t_i = 0, t_{i+1} = 1, t_{i+2} = 2$ ，可以得到五点二次、七点二次和九点二次平滑公式。

五点二次 $\hat{x}_{i-2} = \frac{1}{35} (-3x_{i-2} + 12x_{i-1} + 17x_i + 12x_{i+1} - 3x_{i+2})$ 损失四个数据，无需背诵

七点二次 $\hat{x}_{i-3} = \frac{1}{21} (-2x_{i-3} + 3x_{i-2} + 6x_{i-1} + 7x_i + 6x_{i+1} + 3x_{i+2} - 2x_{i+3})$

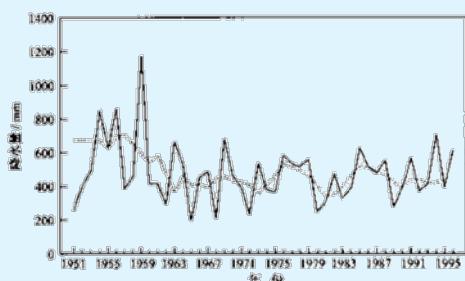
九点二次 $\hat{x}_{i-4} = \frac{1}{231} (-21x_{i-4} + 14x_{i-3} + 39x_{i-2} + 54x_{i-1} + 59x_i + 54x_{i+1} + 39x_{i+2} + 14x_{i+3} - 21x_{i+4})$

5.4.2 计算步骤

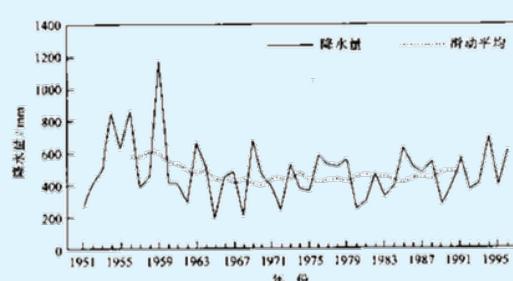
第一步 根据实际问题需要及样本量大小确定平滑的点数 k ，然后根据前面的算式直接对观测数据进行平滑计算，得到**n - k + 1**个平滑值。

第二步 对五、七及九点端点的平滑值，分别由相邻的二、三、四点平滑值平均得到。这样就可以得到n个平滑值。九点便可以保留9年以上的信号，去掉年际的信息。

应用案例



1951-1996 夏季降水九点二次平滑变化趋势
(虚线为平滑值，该图边界存在问题)



1951-1996 夏季降水十一点二次平滑变化趋势
(虚线为平滑值，该图正确)

5.5 变化趋势的显著性检验

引言 对于滑动平均、累积距平等方法得到的变化趋势曲线图往往进行直观判断。对趋势十分明显容易得出结论，而有时则很难直观得到结论，这时可以借助统计检验的办法。
线性倾向估计可以用**相关系数的显著性检验**进行判断。对于**滑动平均、累积距平**等利用**Z检验**。

5.5.1 Z检验方法概述

方法概述 对于气候序列 x_i ，在 i 时刻， $i = 1, 2, \dots, n - 1$ ，有 $r_i = \begin{cases} +1 & \text{当 } x_j > x_i \\ 0 & \text{否则} \end{cases} (j = i + 1, \dots, n)$

可见， r_i 是 i 时刻以后的数值 $x_j, j = i + 1, \dots, n$ 大于该时刻值 x_i 的样本个数。

统计量 $Z = \frac{4\sum_{i=1}^{n-1} r_i}{n(n-1)} - 1$ 显见，对于递增直线， r_i 序列为 $n-1, n-2, \dots, 1$ ，这时 $Z = 1$ ，对于递减直线 $Z = -1$ ， Z 值在 $1 \sim -1$ 之间变化。

检验方法 给定显著性水平 α ，假定 $\alpha = 0.05$ ， $Z_{0.05} = 1.96 \left[\frac{4n+10}{9n(n-1)} \right]^{1/2}$ ，若 $|Z| > Z_{0.05}$ ，则认为变化趋势在 $\alpha = 0.05$ 显著性水平下是显著的。

5.5.2 计算步骤

- 步骤一** 对原气候序列或用某种方法得到的趋势序列，计算其统计量 r_i 。
步骤二 计算统计量 Z 。
步骤三 计算判据 $Z_{0.05}$ ，若 $|Z| > Z_{0.05}$ 则判断变化趋势是显著的。

应用实例

北京 1951—1996 年夏季降水量见表 2.1。用积距平进行变化趋势分析。用上述非参数统计量对变化趋势作显著检验。由方程得到 r_i 序列。 $Z = -0.4531$ ， $Z_{0.05} = 0.20$ ， $|Z| > Z_{0.05}$ ，因此我们认为，在 $\alpha = 0.05$ 显著性水平下，夏季降水量的变化趋势是显著的。

表 4.3 北京夏季降水量累积距平序列的 r_i

12	11	43	29	31	12	13	14	17	9
1	2	1	6	0	1	1	3	3	0
0	3	1	3	5	8	2	1	9	0
0	1	5	8	3	1	1	0	2	4
2	3	3	1	1					